

*Citation for published version:*

Eckhoff, M & Morters, P 2014, 'Vulnerability of robust preferential attachment networks', *Electronic Journal of Probability*, vol. 19, 57, pp. 1-47. <https://doi.org/10.1214/EJP.v19-2974>

*DOI:*

[10.1214/EJP.v19-2974](https://doi.org/10.1214/EJP.v19-2974)

*Publication date:*

2014

*Document Version*

Publisher's PDF, also known as Version of record

[Link to publication](#)

*Publisher Rights*

CC BY

## University of Bath

### Alternative formats

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



Electron. J. Probab. **19** (2014), no. 57, 1–47.  
ISSN: 1083-6489 DOI: 10.1214/EJP.v19-2974

# Vulnerability of robust preferential attachment networks

Maren Eckhoff\*

Peter Mörters\*

## Abstract

Scale-free networks with small power law exponent are known to be robust, meaning that their qualitative topological structure cannot be altered by random removal of even a large proportion of nodes. By contrast, it has been argued in the science literature that such networks are highly vulnerable to a targeted attack, and removing a small number of key nodes in the network will dramatically change the topological structure. Here we analyse a class of preferential attachment networks in the robust regime and prove four main results supporting this claim: After removal of an arbitrarily small proportion  $\epsilon > 0$  of the oldest nodes (1) the asymptotic degree distribution has exponential instead of power law tails; (2) the largest degree in the network drops from being of the order of a power of the network size  $n$  to being just logarithmic in  $n$ ; (3) the typical distances in the network increase from order  $\log \log n$  to order  $\log n$ ; and (4) the network becomes vulnerable to random removal of nodes. Importantly, all our results explicitly quantify the dependence on the proportion  $\epsilon$  of removed vertices. For example, we show that the critical proportion of nodes that have to be retained for survival of the giant component undergoes a steep increase as  $\epsilon$  moves away from zero, and a comparison of this result with similar ones for other networks reveals the existence of two different universality classes of robust network models. The key technique in our proofs is a local approximation of the network by a branching random walk with two killing boundaries, and an understanding of the particle genealogies in this process, which enters into estimates for the spectral radius of an associated operator.

**Keywords:** Power law; small world; scale-free network; preferential attachment; Barabási-Albert model; percolation; maximal degree; diameter; network distance; robustness; vulnerability; multitype branching process; killed branching random walk.

**AMS MSC 2010:** Primary 05C80, Secondary 60J85; 60K35; 90B15.

Submitted to EJP on August 22, 2013, final version accepted on June 24, 2014.

Supersedes arXiv:1306.3454v2.

## 1 Introduction

### 1.1 Motivation

The problem of resilience of networks to either random or targeted attack is crucial to many instances of real world networks, ranging from social networks (like collaboration networks) via technological networks (like electrical power grids) to communication networks (like the world-wide web). Of particular importance is whether

\*University of Bath, United Kingdom. E-mail: eckhoff.maren@gmail.com, maspm@bath.ac.uk

the connectivity of a network relies on a small number of hubs and whether their loss will cause a large-scale breakdown. Albert, Albert and Nakarado [1] argue that “the power grid is robust to most perturbations, yet disturbances affecting key transmission substations greatly reduce its ability to function”. Experiments of Albert, Jeong, and Barabási [2], Holme, Kim, Yoon and Han [24] and more recently of Mishkovski, Biey and Kocarev [29] find robustness under random attack but vulnerability to the removal of a small number of key nodes in several other networks. The latter paper includes a study of data related to the human brain, as well as street, collaboration and power grid networks. One should expect this qualitative behaviour across the range of real world networks and it should therefore also be present in the key mathematical models of large complex networks.

A well established feature of many real world networks is that in a suitable range of values  $k$  the proportion of nodes with degree  $k$  has a decay of order  $k^{-\tau}$  for a power law exponent  $\tau$ . The robustness of networks with small power law exponent under random attack has been observed heuristically by Callaway et al. [8] and Cohen et al. [11], but there seems to be controversy in these early papers about the extent of the vulnerability in the case of targeted attack, see the discussion in [17] and [12]. As Bollobás and Riordan [7, Section 10] point out, such heuristics, informative as they may be, are often quite far away from a mathematical proof that applies to a given model. In their seminal paper [7] they provide the first rigorous proof of robustness in the case of a specific preferential attachment model with power law exponent  $\tau = 3$ , and later Dereich and Mörters [15] proved for a class of preferential attachment models with tunable power law exponent that networks are robust under random attack if the power law exponent satisfies  $\tau \leq 3$ , but not when  $\tau > 3$ , thus revealing the precise location of the phase transition in the behaviour of preferential attachment networks. However, the question of vulnerability of robust networks when a small number of privileged nodes is removed has not been studied systematically in the mathematical literature so far.

It is the aim of the present paper to give evidence for the vulnerability of robust networks by providing rigorous proof that preferential attachment networks in the robust regime  $\tau \leq 3$  undergo a radical change under a targeted attack, i.e. when an arbitrarily small proportion  $\epsilon > 0$  of the most influential nodes in the network is removed. Our main results, presented in Section 1.3, show how precisely this change affects the degree structure, the length of shortest paths and the connectivity in the network. The results take the form of limit theorems revealing explicitly the dependence of the relevant parameters on  $\epsilon$ . Not only does this provide further insight into the topology of the network and the behaviour as  $\epsilon$  tends to zero, it also allows a comparison to other network models, and thus exposes two classes of robust networks with rather different behaviour, see Section 1.5. Our mathematical analysis of the network uses several new ideas and combines probabilistic and combinatorial arguments with analytic techniques informed by new probabilistic insights. It is crucially based on the local approximation of preferential attachment networks by a branching random walk with a killing boundary recently found in [15]. In this approximation the removal of a proportion of old vertices corresponds to the introduction of a second killing boundary. On the one hand this adds an additional level of complexity to the process, as the mathematical understanding of critical phenomena in branching models on finite intervals is only just emerging, see for example [20]. On the other hand compactness of the typespace for this branching process opens up new avenues that are exploited, for example, in the form of spectral estimates based on rather subtle information on the shape of principal eigenfunctions of an operator associated with the branching process.

## 1.2 Mathematical framework

The established mathematical model for a large network is a sequence  $(G_n: n \in \mathbb{N})$  of (random or deterministic) graphs  $G_n$  with vertex set  $V_n$  and an edge set  $E_n$  consisting of (directed or undirected) edges between the vertices. We assume that the size  $|V_n|$  of the vertex set is increasing to infinity in probability, so that results about the limiting behaviour in the sequence of graphs may be seen as predictions for the behaviour of large networks. In all cases of interest here the average number of edges per vertex converges in probability to a finite limit and the topology of a bounded neighbourhood of a typical vertex stabilizes. An important example for this is the proportion of vertices with a given degree in  $G_n$ , which in the relevant models converges and allows us to talk about the asymptotic degree distribution. The mathematical models of power law networks therefore have an asymptotic degree distribution with the probability of degree  $k$  decaying like  $k^{-\tau}$ , as  $k \rightarrow \infty$ , for some  $\tau > 1$ . Our focus here is on the global properties emerging in network models with asymptotic power law degree distributions.

A crucial global feature of a network is its connectivity, and in particular the existence of a large connected component. To describe this, we denote by  $C_n$  a connected component in  $G_n$  with maximal number of nodes. The graph sequence  $(G_n: n \in \mathbb{N})$  has a *giant component* if there exists a constant  $\zeta > 0$  such that

$$\frac{|C_n|}{\mathbb{E}|V_n|} \rightarrow \zeta \quad \text{as } n \rightarrow \infty,$$

where the convergence holds in probability. We remark that for the models usually considered the issue is not the convergence itself but the positivity of the limit  $\zeta$ . If a giant component exists and the length of the shortest path between any two vertices in the largest component of  $G_n$  is asymptotically bounded by a multiple of  $\log n$ , the network is called *small*. If it is asymptotically bounded by a constant multiple of  $\log \log n$ , the network is called *ultrasmall*; cf. Section 1.2 in [23].

To model a *random attack* on the network we keep each vertex in  $G_n$  independently with probability  $p \in [0, 1]$  and otherwise we remove it from the vertex set together with all its adjacent edges, i.e. we run *vertex percolation* on  $G_n$  with *retention probability*  $p$ . The resulting graph is denoted by  $G_n(p)$ . A simple coupling argument shows that there exists a critical parameter  $p_c \in [0, 1]$  such that the sequence  $(G_n(p): n \in \mathbb{N})$  has a giant component if  $p_c < p \leq 1$ , and it does not have a giant component if  $0 \leq p < p_c$ . If  $p_c = 0$ , i.e. if the giant component cannot be destroyed by percolation with any positive retention parameter, then the network is called *robust*. To study the resilience of networks to a *targeted attack* we look at models in which the construction of the network favours certain vertices in such a way that these privileged vertices have a better chance of getting a high degree than others. When  $G_n$  is a network on  $n$  vertices, we label these by 1 to  $n$  and assume that vertices are ordered in decreasing order of privilege. This assumption allows an attacker to target the most privileged vertices without knowledge of the entire graph. The *damaged graph*  $G_n^\epsilon$ , for some  $\epsilon \in (0, 1)$ , is obtained from  $G_n$  by the removal of all vertices with label less or equal to  $\epsilon n$  together with all adjacent edges. In particular, the new vertex set is  $V_n^\epsilon = \{\lfloor \epsilon n \rfloor + 1, \dots, n\}$ , and we let  $C_n^\epsilon$  be a connected component in  $G_n^\epsilon$  with maximal number of nodes. We write  $G_n^\epsilon(p)$  for the graph obtained from  $G_n$  by first removing all vertices with label at most  $\epsilon n$  and then running vertex percolation on the remaining graph. Note that we would get the same graph when reversing the order in which these two attacks are performed. However, we always start with the targeted attack for definiteness.

We investigate the problem of vulnerability of random networks to targeted attacks in the context of *preferential attachment networks*. This class of models has been pop-

ularised by Barabási and Albert [3] and has received considerable attention in the scientific literature. The idea is that a sequence of graphs is constructed by successively adding vertices. Together with a new vertex, edges are introduced by connecting it to existing vertices at random with a probability depending on the degree of the existing node; the higher the degree the more likely the connection. Despite the relatively simple principle on which this model is based it shows a good match of global features with real networks. For example, the asymptotic degree distributions follow a power law, and variations in the attachment probabilities allow for tuning of the power law exponent  $\tau$ . If the power law exponent satisfies  $\tau < 3$ , then the network is robust and ultrasmall.

The first mathematically rigorous study of resilience in preferential attachment networks was performed by Bollobás and Riordan [7] for the so-called LCD model. This model variant has the advantage of having an explicit static description, which makes it easier to analyse than models that have only a dynamic description. It also has a fixed power law exponent  $\tau = 3$ , hence, Bollobás and Riordan [7] prove only results for this specific exponent. They show that in this case the network is robust and identify a critical proportion  $\epsilon_c < 1$  such that the removal of the oldest  $\lfloor \epsilon n \rfloor$  oldest vertices leads to the destruction of the giant component if and only if  $\epsilon \geq \epsilon_c$ . Note that this is not in line with the notion of vulnerability that we are interested in as we only want to remove a *small* proportion of old vertices.

In the present paper, we consider the question of vulnerability in the following model variant, introduced in [14]. Let  $\mathbb{N}_0$  be the set of nonnegative integers and fix a function  $f: \mathbb{N}_0 \rightarrow (0, \infty)$ , which we call the *attachment rule*. The most important case is if  $f$  is affine, i.e.  $f(k) = \gamma k + \beta$  for parameters  $\gamma \in [0, 1)$  and  $\beta > 0$ , but non-linear functions are allowed. Given an attachment rule  $f$ , we define a growing sequence  $(G_n: n \in \mathbb{N})$  of random graphs by the following dynamics:

- Start with one vertex labelled 1 and no edges, i.e.  $G_1$  is given by  $V_1 := \{1\}$ ,  $E_1 := \emptyset$ ;
- Given the graph  $G_n$ , we construct  $G_{n+1}$  from  $G_n$  by adding a new vertex labelled  $n+1$  and, for each  $m \leq n$  independently, inserting the directed edge  $(n+1, m)$  with probability

$$\frac{f(\text{indegree of } m \text{ at time } n)}{n} \wedge 1. \quad (1.1)$$

Formally we are dealing with a sequence of directed graphs but all edges point from the younger to the older vertex. Hence, the directions can be recreated from the undirected, labelled graph. For all structural questions, particularly regarding connectivity and the length of shortest paths, we regard  $(G_n: n \in \mathbb{N})$  as an undirected network. Dereich and Mörters consider in [14, 15] concave attachment rules  $f$ . Denoting the asymptotic slope of  $f$  by

$$\gamma := \lim_{k \rightarrow \infty} \frac{f(k)}{k}, \quad (1.2)$$

they show that for  $\gamma \in (0, 1)$  the sequence  $(G_n: n \in \mathbb{N})$  has an asymptotic degree distribution which follows a power law with exponent

$$\tau = \frac{\gamma + 1}{\gamma}.$$

For  $\gamma \geq 1$ , i.e.  $\tau \leq 2$ , the mean of the asymptotic degree distribution is infinite and a radically different topology can be expected. Results on power law networks in this regime have been derived for example in [18, 4]; we restrict ourselves to the finite mean case  $\gamma < 1$ . In the case  $\gamma < \frac{1}{2}$ , or equivalently  $\tau > 3$ , there exists a critical percolation

parameter  $p_c > 0$  such that  $(G_n(p): n \in \mathbb{N})$  has a giant component if and only if  $p > p_c$ .<sup>1</sup> If however  $\gamma \geq \frac{1}{2}$ , or equivalently  $\tau \leq 3$ , the sequence  $(G_n(p): n \in \mathbb{N})$  has a giant component for all  $p \in (0, 1]$ , i.e.  $(G_n: n \in \mathbb{N})$  is robust. This is the regime of interest in the present paper.

### 1.3 Statement of the main results

Our focus is on the case of an affine attachment rule  $f(k) = \gamma k + \beta$  with  $\beta > 0$  and  $\gamma \in [\frac{1}{2}, 1)$ . Recall that for this choice the preferential attachment network is robust. We use the symbol  $a(\epsilon) \asymp b(\epsilon)$  to indicate that there are constants  $0 < c < C$  and some  $\epsilon_0 > 0$  such that  $cb(\epsilon) \leq a(\epsilon) \leq Cb(\epsilon)$  for all  $0 < \epsilon < \epsilon_0$ .

**Theorem 1.1. (Loss of connectivity)** *For any  $\epsilon \in (0, 1)$ , there exists  $p_c(\epsilon) \in (0, 1]$  such that the damaged network*

$$(G_n^\epsilon(p): n \in \mathbb{N}) \text{ has a giant component} \Leftrightarrow p > p_c(\epsilon). \quad (1.3)$$

If  $\gamma = \frac{1}{2}$  then

$$p_c(\epsilon) \asymp \frac{1}{\log(1/\epsilon)}. \quad (1.4)$$

If  $\gamma > \frac{1}{2}$  then, as  $\epsilon \downarrow 0$ ,

$$p_c(\epsilon) = \frac{2\gamma - 1}{\sqrt{\beta(\gamma + \beta)}} \epsilon^{\gamma-1/2} [1 + O(\epsilon^{\gamma-1/2}(\log \epsilon))].$$

Theorem 1.1 shows that the removal of an arbitrarily small proportion of old nodes makes the network vulnerable to percolation, but does not destroy the giant component. The steep increase of  $p_c(\epsilon)$  as  $\epsilon$  leaves zero shows that, even when a small proportion of old nodes has been removed in the network, the removal of further old nodes is much more destructive than the removal of a similar proportion of randomly chosen nodes.

As for small  $\epsilon$  the critical value  $p_c(\epsilon)$  is strictly decreasing in  $\gamma$ , this effect is stronger the closer  $\gamma$  is to  $\frac{1}{2}$ . This result might be perceived as slightly counterintuitive since the preferential attachment becomes stronger as  $\gamma$  increases and therefore we might expect older nodes to be more privileged and a targeted attack to be more effective than in the small  $\gamma$  regime. However, the effect of the stronger preferential attachment is more than compensated by the fact that networks with a small value of  $\gamma$  have a (stochastically) smaller number of edges and are therefore a-priori more vulnerable. Note also that  $p_c(\epsilon)$  may be equal to 1 if  $\epsilon$  is not sufficiently small in which case (1.3) implies that the damaged network has no giant component. In the case  $\gamma = \frac{1}{2}$  the implied constants in (1.4) can be made explicit as  $c = \frac{1}{\gamma + \beta}$  and  $C = \frac{1}{\beta}$ , but we cannot show that they match asymptotically. However, we conjecture that

$$p_c(\epsilon) \sim \frac{1}{\sqrt{\beta(\gamma + \beta)}} \frac{1}{\log(1/\epsilon)} \quad \text{as } \epsilon \downarrow 0,$$

meaning that the ratio of the left- and right-hand sides converges to one.

To gain further insight into the topology of the damaged graph, we now look at the asymptotic indegree distribution and at the largest indegree in the network. Recall from [14] that outdegrees are asymptotically Poisson distributed and therefore indegrees are solely responsible for the power law behaviour as well as the dynamics of maximal degrees. From here onwards we additionally assume that  $\beta \leq 1$ . Under this

<sup>1</sup>The results of [15] are formulated for edge percolation, whereas we consider vertex percolation. It is not hard to see that for the existence or nonexistence of the giant component this makes no difference.

condition,  $f(n) < n + 1$  for all  $n \in \mathbb{N}_0$  and the minimum in (1.1) is always attained by its first argument.

For a probability measure  $\nu$  on the nonnegative integers, we write  $\nu_{\geq k} := \nu(\{k, k + 1, \dots\})$  and  $\nu_k := \nu(\{k\})$ . Let  $\mathcal{Z}[m, n]$  be the indegree of vertex  $m$  in  $G_n$  at time  $n \geq m$ . Since for  $m > \lfloor \epsilon n \rfloor$ , the indegree of  $m$  in  $G_n$  and  $G_n^\epsilon$  agree, writing  $X^\epsilon(n)$  for the empirical indegree distribution in  $G_n^\epsilon$ , we have

$$X_k^\epsilon(n) = \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{m=\lfloor \epsilon n \rfloor+1}^n \mathbb{1}_{\{k\}}(\mathcal{Z}[m, n]), \quad \text{for } k \in \mathbb{N}_0.$$

We write  $M(G)$  for the maximal indegree in a directed graph  $G$ . For  $s, t > 0$ , let

$$B(s, t) := \int_0^1 x^{s-1} (1-x)^{t-1} dx$$

denote the beta function at  $(s, t)$ . Before we make statements about the network after the targeted attack, we recall the situation in the undamaged network. In this case Dereich and Mörters in [14] show that the empirical indegree distribution  $X^0(n)$  in  $G_n$  satisfies almost surely

$$\lim_{n \rightarrow \infty} X^0(n) = \mu$$

in total variation norm. The limit is the probability measure  $\mu$  on the nonnegative integers given by

$$\mu_{\geq k} = \frac{B(k + \frac{\beta}{\gamma}, \frac{1}{\gamma})}{B(\frac{\beta}{\gamma}, \frac{1}{\gamma})} \quad \text{for } k \in \mathbb{N}_0,$$

and satisfies  $\lim_{k \rightarrow \infty} \log \mu_{\geq k} / \log k = -1/\gamma$ . Moreover, the maximal indegree satisfies, in probability,

$$\frac{\log M(G_n)}{\log(n^\gamma)} \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

Our theorem shows that in the damaged network the asymptotic degree distribution is no longer a power law but has exponential tails. The maximal degree grows only logarithmically, not polynomially.

**Theorem 1.2. (Collapse of large degrees)** *Let  $\epsilon \in (0, 1)$ . Almost surely,*

$$\lim_{n \rightarrow \infty} X^\epsilon(n) = \mu^\epsilon$$

*in total variation norm. The limit is the probability measure  $\mu^\epsilon$  on the nonnegative integers given by*

$$\mu_{\geq k}^\epsilon = \int_\epsilon^1 \frac{1}{1-\epsilon} B(k, \frac{\beta}{\gamma})^{-1} \int_{y\gamma}^1 x^{\frac{\beta}{\gamma}-1} (1-x)^{k-1} dx dy \quad \text{for } k \in \mathbb{N}_0.$$

*It satisfies  $\lim_{k \rightarrow \infty} \log \mu_{\geq k}^\epsilon / k = \log(1 - \epsilon^\gamma)$ . Moreover, the maximal indegree satisfies, in probability,*

$$\frac{M(G_n^\epsilon)}{\log n} \rightarrow -\frac{1}{\log(1 - \epsilon^\gamma)} \quad \text{as } n \rightarrow \infty. \quad (1.5)$$

It is worth mentioning that  $\mu = \mu^0$ , so Theorem 1.2 remains valid for  $\epsilon = 0$ . Moreover, the result holds also for  $\gamma \in (0, \frac{1}{2})$  by the same proof. Theorem 1.2 shows in particular that by removing a proportion of the oldest vertices we have removed all vertices with a degree bigger than a given constant multiple of  $\log n$ . This justifies the comparison of our vulnerability results with empirical studies of real world networks such as [1], in which all nodes whose degree exceeds a given threshold are removed. Note also that as  $\epsilon \downarrow 0$  the right-hand side in (1.5) is asymptotically equivalent to  $\epsilon^{-\gamma}$  and the growth of the maximal degree is the faster the larger  $\gamma$ .

Preferential attachment networks are ultrasmall for sufficiently small power law exponents. For our model, Mönch [31], see also [13, 16], has shown that, denoting by  $d_G$  the graph distance in a graph  $G$ , for independent random vertices  $V_n, W_n$  chosen uniformly from  $C_n$ , we have

$$\begin{aligned} \text{if } \gamma = \frac{1}{2}, \text{ then} & \quad d_{G_n}(V_n, W_n) \sim \frac{\log n}{\log \log n}, \\ \text{if } \gamma > \frac{1}{2}, \text{ then} & \quad d_{G_n}(V_n, W_n) \sim \frac{4 \log \log n}{\log(\gamma/(1-\gamma))}, \end{aligned}$$

meaning that the ratio of the left- and right-hand side converges to one in probability as  $n \rightarrow \infty$ . Removing an arbitrarily small proportion of old vertices however leads to a massive increase in the typical distances, as our third main theorem reveals. We say that a sequence of events  $(\mathcal{E}_n: n \in \mathbb{N})$  holds *with high probability* if  $\mathbb{P}(\mathcal{E}_n) \rightarrow 1$  as  $n \rightarrow \infty$ .

**Theorem 1.3. (Increase of typical distances)** *Let  $\epsilon > 0$  be sufficiently small so that  $(G_n^\epsilon: n \in \mathbb{N})$  has a giant component, and let  $V_n, W_n$  be chosen independently and uniformly from  $C_n^\epsilon$ . Then, for all  $\delta > 0$ ,*

$$d_{G_n^\epsilon}(V_n, W_n) \geq \frac{1-\delta}{\log(1/p_c(\epsilon))} \log n \quad \text{with high probability.}$$

Our proof gives the result for all values  $\gamma \in [0, 1)$ ,  $\epsilon > 0$ , with  $p_c(\epsilon) < 1$  but if  $\gamma < \frac{1}{2}$  even without removal of old vertices the typical distances in the network are known to be of order  $\log n$ , so that this is not surprising. We believe that there is an upper bound matching the lower bound above, but the proof would be technical and the result much less interesting.

In the next two sections we discuss some further ramifications of our main results.

#### 1.4 Non-linear attachment rules

So far we have presented results for the case of affine attachment rules  $f$ , given by  $f(k) = \gamma k + \beta$ . While the fine details of the network behaviour often depend on the exact model definition, we expect the principal scaling and macroscopic features to be independent of these details. To investigate this *universality* we now discuss to what extent Theorem 1.1 remains true when we look at more general non-linear attachment rules  $f$ .

We consider two classes of attachment rules.

- (1) A function  $f: \mathbb{N}_0 \rightarrow (0, \infty)$  is called a *L-class* attachment rule if there exists  $\gamma \in [0, 1)$  and  $0 < \beta_l \leq \beta_u$  such that  $\gamma k + \beta_l \leq f(k) \leq \gamma k + \beta_u$  for all  $k$ . Note that the parameter  $\gamma$  for a L-class rule is uniquely defined by (1.2).
- (2) A concave function  $f: \mathbb{N}_0 \rightarrow (0, \infty)$  with  $\gamma := \lim_{k \rightarrow \infty} f(k)/k \in [0, 1)$  is called a *C-class* attachment rule. Note that concavity of  $f$  implies that the limit above exists and that  $f$  is non-decreasing.

The asymptotic slope of the attachment rule determines the key features of the model. For example, Dereich and Mörters [14] show that for certain *C-class* attachment rules with  $\gamma > 0$  the asymptotic degree distribution is a power law with exponent  $\tau = 1 + 1/\gamma$ . The following theorem shows that  $\gamma$  also determines the scaling of the critical percolation probability for the damaged network.



**Theorem 1.4. (Loss of connectivity, non-linear case)** *Let  $f$  be a L-class or C-class attachment rule. For all  $\epsilon \in (0, 1)$ ,*

$$p_c(\epsilon) := \inf \{p: (G_n^\epsilon(p): n \in \mathbb{N}) \text{ has a giant component}\} > 0.$$

Moreover, if  $f$  is in the L-class and

$$\begin{aligned} \text{if } \gamma = \frac{1}{2}, \text{ then} & \quad \lim_{\epsilon \downarrow 0} \frac{\log p_c(\epsilon)}{\log \log(1/\epsilon)} = -1, \\ \text{if } \gamma > \frac{1}{2}, \text{ then} & \quad \lim_{\epsilon \downarrow 0} \frac{\log p_c(\epsilon)}{\log \epsilon} = \gamma - \frac{1}{2}. \end{aligned}$$

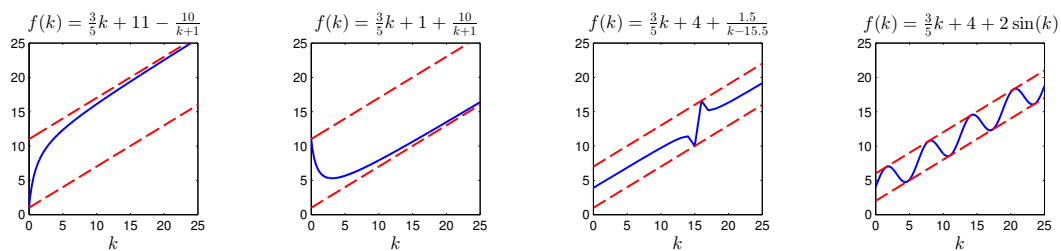
If  $f$  is in the C-class, the statement remains true in the case  $\gamma > \frac{1}{2}$ , and in the case  $\gamma = \frac{1}{2}$  if the limit is replaced by a  $\limsup_{\epsilon \downarrow 0}$  and the equality by ' $\leq$ '.

Theorem 1.4 implies that the damaged network  $(G_n^\epsilon: n \in \mathbb{N})$  is not robust. But as  $\lim_{\epsilon \downarrow 0} p_c(\epsilon) = 0$  it is still 'asymptotically robust' for  $\epsilon \downarrow 0$  in the sense that when less than order  $n$  old vertices are destroyed, then the critical percolation parameter remains zero. We formulate this as a corollary. For two graphs  $G = (V, E)$  and  $\tilde{G} = (\tilde{V}, \tilde{E})$ , we write  $G \geq \tilde{G}$  if there is a coupling such that  $V \supseteq \tilde{V}$  and  $E \supseteq \tilde{E}$ .

**Corollary 1.5.** *Let  $f$  be a L-class or C-class attachment rule with  $\gamma \geq \frac{1}{2}$  and let  $(m_n: n \in \mathbb{N})$  be a sequence of natural numbers with  $\lim_{n \rightarrow \infty} m_n/n = 0$ . The network  $(G_n^{(m_n)}: n \in \mathbb{N})$ , consisting of the graphs  $G_n$  damaged by removal of the oldest  $m_n$  vertices along with all adjacent edges, is robust.*

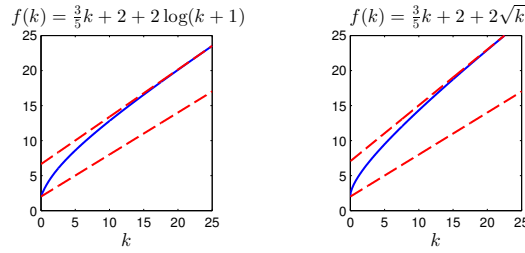
*Proof.* Let  $p \in (0, 1)$ . By Theorem 1.4, there exists  $\epsilon > 0$  such that  $p_c(\epsilon) < p$ . Choose  $n_0 \in \mathbb{N}$  such that  $m_n/n < \epsilon$  for all  $n \geq n_0$ . Then  $G_n^{(m_n)} \geq G_n^\epsilon$  for all  $n \geq n_0$ , implying  $G_n^{(m_n)}(p) \geq G_n^\epsilon(p)$ . Since the network  $(G_n^\epsilon(p): n \in \mathbb{N})$  has a giant component, so does the network  $(G_n^{(m_n)}(p): n \in \mathbb{N})$ .  $\square$

Theorem 1.4 is derived from Theorem 1.1 using the monotonicity of the network in the attachment rule. Its appeal lies in the large class of functions to which it applies. The L-class attachment rules are all positive, bounded perturbations of linear functions. In Figure 1 we see several examples: On the left a concave function which is also in the C-class, then a convex function and a function which is convex in one and concave in another part of its domain. The latter examples are not monotone, and all three are asymptotically vanishing perturbations of an affine attachment rule. The example of an L-class attachment rule on the right shows that this may also fail.



**Figure 1:** Examples for L-class attachment rules. The blue curve is the attachment rule, the red, dashed lines are linear lower and upper bounds.

The C-class attachment rules are always non-decreasing as positive concave functions and always have a linear lower bound with the same asymptotic slope  $\gamma$  as the function itself. However, when the perturbation  $k \mapsto f(k) - \gamma k$  is not bounded, then there exists no linear function with slope  $\gamma$  which is an upper bound to the attachment rule; any linear upper bound will be steeper. Two examples are displayed in Figure 2.



**Figure 2:** Examples for C-class attachment rules. The blue curve is the attachment rule, the red, dashed lines are linear lower and upper bounds. The slope of the upper bound is strictly larger than  $\gamma$ .

### 1.5 Vulnerability of other network models

We would like to investigate to what extent our results are common to robust random network models rather than specific to preferential attachment networks. Again our focus is on Theorem 1.1 and we look at two types of networks, the configuration model and the inhomogeneous random graphs. Both types have an explicit static description and are therefore much easier to analyse than the preferential attachment networks studied in our main theorems.

#### 1.5.1 Configuration model

A targeted attack can be planned particularly well when the degree sequence of the network is known. A random graph model with fixed degree sequence is given by the configuration model. For  $n \in \mathbb{N}$ , let  $\mathbf{d}_n = (d_i^n)_{i=1}^n \in \mathbb{N}^n$  with  $\sum_{i=1}^n d_i^n$  even. To simplify notation, we write  $d_i$  instead of  $d_i^n$ . The multigraph  $G_n^{(\text{CM})}$  on vertex set  $\{1, \dots, n\}$  is constructed as follows: to every vertex  $i$  attach  $d_i$  half-edges. Combine the half-edges into pairs by a uniformly random matching of the set of all half-edges. Each pair of half-edges is then joined to form an edge of  $G_n^{(\text{CM})}$ . The configuration model received a lot of attention in the literature, see [23] and the references therein. A good targeted attack in the configuration model is the removal of the vertices with the highest degree and we denote by  $G_n^{(\text{CM}), \epsilon}$  the network after removal of the  $\lfloor \epsilon n \rfloor$  vertices with the largest degree.

Let  $n_k = |\{i \leq n : d_i = k\}|$  be the number of vertices with degree  $k$  and assume that there exists a  $\mathbb{N}$ -valued random variable  $D$  with  $0 < \mathbb{E}D < \infty$  and  $\mathbb{P}(D = 2) < 1$ , such that

$$n_k/n \rightarrow \mathbb{P}(D = k) \quad \text{for all } k \in \mathbb{N} \quad \text{and} \quad \frac{1}{n} \sum_{k=1}^{\infty} k n_k \rightarrow \mathbb{E}D \quad \text{as } n \rightarrow \infty. \quad (1.6)$$

In particular, the law of  $D$  is the weak limit of the empirical degree distribution in  $(G_n^{(\text{CM})} : n \in \mathbb{N})$  and the network is robust if  $\mathbb{E}[D^2] = \infty$ ; see [25, Theorem 3.5]. Our focus is on the case that the distribution of  $D$  is a power law with exponent  $\tau = 1 + 1/\gamma$ ,  $\gamma \geq \frac{1}{2}$ .

**Theorem 1.6.** *Let  $\epsilon \in (0, 1)$ ,  $\gamma \geq \frac{1}{2}$  and suppose there is a constant  $C > 0$  such  $\mathbb{P}(D > k) \sim Ck^{-1/\gamma}$  as  $k \rightarrow \infty$ . Then there exists  $p_c^{(\text{CM})}(\epsilon) > 0$  such that*

$$(G_n^{(\text{CM}), \epsilon}(p) : n \in \mathbb{N}) \text{ has a giant component} \Leftrightarrow p > p_c^{(\text{CM})}(\epsilon).$$

Moreover,

$$p_c^{(\text{CM})}(\epsilon) \asymp \begin{cases} \frac{1}{\log(1/\epsilon)} & \text{if } \gamma = \frac{1}{2}, \\ \epsilon^{2\gamma-1} & \text{if } \gamma > \frac{1}{2}. \end{cases}$$

We observe the same basic phenomenon as in the corresponding preferential attachment models: While the undamaged network is robust, after removal of an arbitrarily small proportion of privileged nodes the network becomes vulnerable to random removal of vertices. However, when  $\gamma > \frac{1}{2}$ , then the increase of the critical percolation parameter  $p_c(\epsilon)$  as  $\epsilon$  leaves zero is less steep than in the corresponding preferential attachment model.

Note that our assumptions imply that  $0 < \mathbb{E}D < \infty$ . In the case  $\mathbb{E}D = \infty$ , Bhamidi et al. [4] show a more extreme form of vulnerability, where the connected network can be disconnected with high probability by deleting a bounded number of vertices.

### 1.5.2 Inhomogeneous random graphs

Inhomogeneous random graphs are a generalization of the classical Erdős-Rényi random graph. Let  $\kappa: (0, 1] \times (0, 1] \rightarrow (0, \infty)$  be a symmetric kernel. The inhomogeneous random graph  $G_n^{(\kappa)}$  corresponding to kernel  $\kappa$  has the vertex set  $V_n = \{1, \dots, n\}$  and any pair of distinct vertices  $i$  and  $j$  is connected by an edge independently with probability

$$\mathbb{P}(\{i, j\} \text{ present in } G_n^{(\kappa)}) = \frac{1}{n} \kappa\left(\frac{i}{n}, \frac{j}{n}\right) \wedge 1. \quad (1.7)$$

Many features of this class of models are discussed by Bollobás, Janson and Riordan [6] and van der Hofstad [23]. The first inhomogeneous random graph model we consider is a version of the *Chung-Lu model*, see for example [9, 10]. The relevant kernel is

$$\kappa^{(\text{CL})}(x, y) = x^{-\gamma} y^{-\gamma} \quad \text{for } x, y \in (0, 1].$$

This is an example of a kernel of the form  $\kappa(x, y) = \psi(x)\psi(y)$ , for some  $\psi$ , which are called kernels of *rank one*, see [6]. Note that a similar factorisation occurs in the configuration model since the probability that vertices  $i$  and  $j$  are directly connected is roughly proportional to  $d_i d_j$ . Therefore, the configuration model can be classified as a *rank one model*, too. The network corresponding to  $\kappa^{(\text{CL})}$  has an asymptotic degree distribution which is a power law with exponent  $\tau = 1 + 1/\gamma$ .

The second inhomogeneous random graph model we consider is chosen such that the edge probabilities agree (at least asymptotically) with those in a preferential attachment network and the asymptotic degree distribution is a power law with exponent  $\tau = 1 + 1/\gamma$ . The relevant kernel is

$$\kappa^{(\text{PA})}(x, y) = \frac{1}{(x \wedge y)^\gamma (x \vee y)^{1-\gamma}} \quad \text{for } x, y \in (0, 1].$$

Note that if  $\gamma \neq \frac{1}{2}$  this kernel is not of rank one, but *strongly inhomogeneous*. These two kernels allow us to demonstrate the difference between rank one models and preferential attachment models within one model class.

We denote by  $G_n^{(\text{CL})}$ , resp.  $G_n^{(\text{PA})}$  the inhomogeneous random graphs with kernel  $\kappa^{(\text{CL})}$ , resp. kernel  $\kappa^{(\text{PA})}$ . If  $\gamma \geq \frac{1}{2}$ , then  $(G_n^{(\text{CL})} : n \in \mathbb{N})$  and  $(G_n^{(\text{PA})} : n \in \mathbb{N})$  are robust. Since the kernels  $\kappa^{(\text{CL})}$  and  $\kappa^{(\text{PA})}$  are decreasing in both components, vertices with small labels are favoured in the corresponding models. We denote by  $G_n^{(\text{CL}), \epsilon}$ , resp.  $G_n^{(\text{PA}), \epsilon}$ , what remains of the graph  $G_n^{(\text{CL})}$ , resp.  $G_n^{(\text{PA})}$ , after removal of all vertices with label at most  $\epsilon n$  along with their adjacent edges.

The following theorem confirms that, like in the preferential attachment and in the configuration model, the removal of a positive fraction of key vertices makes the networks vulnerable to random removal of nodes. Notice that  $\kappa^{(\text{CL})}$  and  $\kappa^{(\text{PA})}$  agree for  $\gamma = \frac{1}{2}$  so that we only have to state a result for  $G_n^{(\text{CL}), \epsilon}$  in this regime.

**Theorem 1.7.** *Let  $\gamma \geq \frac{1}{2}$ ,  $\star \in \{\text{CL}, \text{PA}\}$  and  $\epsilon \in (0, 1)$ . There exists  $p_c^{(\star)}(\epsilon) > 0$  such that*

$$(G_n^{(\star), \epsilon}(p) : n \in \mathbb{N}) \text{ has a giant component} \Leftrightarrow p > p_c^{(\star)}(\epsilon).$$

Moreover,

$$p_c^{(\text{CL})}(\epsilon) = \begin{cases} \frac{1}{\log(1/\epsilon)} & \text{if } \gamma = \frac{1}{2}, \\ (2\gamma - 1)\epsilon^{2\gamma-1}[1 + O(\epsilon^{2\gamma-1})] & \text{if } \gamma > \frac{1}{2}, \end{cases}$$

and

$$p_c^{(\text{PA})}(\epsilon) \asymp \epsilon^{\gamma-1/2} \quad \text{if } \gamma > \frac{1}{2}.$$

The fact that the Chung-Lu model is vulnerable to targeted attacks has also been remarked by van der Hofstad in Section 9.1 of [23].

Summarising, we note that vulnerability to a targeted attack is a universal feature of robust networks, holding not only for preferential attachment networks but also for configuration models and various classes of inhomogeneous random graphs. In the case  $2 < \tau < 3$ , studying the asymptotic behaviour of the critical percolation parameter  $p_c(\epsilon)$  as a function of the proportion  $\epsilon$  of removed vertices reveals *two universality classes* of networks, distinguished by the critical exponent measuring the polynomial rate of decay of  $p_c(\epsilon)$  as  $\epsilon \downarrow 0$ . In terms of the power law exponent  $\tau$  this critical exponent equals  $\frac{3-\tau}{\tau-1}$  in the case of the configuration model and the Chung-Lu model, but is only half this value in the case of preferential attachment networks and inhomogeneous random graphs with a strongly inhomogeneous kernel. The same classification of networks has emerged in a different context in [13], where it was noted that the typical distances in networks of the two classes differ by a factor of two. The key feature of the configuration model and the rank one inhomogeneous random graphs seems to be that the connection probability of two vertices factorises. By contrast, the connection probabilities in preferential attachment networks have a more complex structure giving privileged nodes a stronger advantage.

## 1.6 The local neighbourhood in the network

Dereich and Mörters [15] have shown that the (not too large) graph neighbourhood of a uniformly chosen vertex in  $G_n$  can be coupled to a branching random walk on the negative half-line. Although we cannot make direct use of this coupling result in our proofs, it is helpful to formulate our ideas in this framework. Therefore, we now explain heuristically that a suitable exploration of the local neighbourhood of a given vertex  $v_0 \in G_n^\epsilon$  reveals a graph that can be approximated by the genealogical tree of a two-type branching random walk with two killing boundaries. A complete definition of the branching process used in the current article is given in Section 2.1 and the coupling is proved rigorously in Section 4 below.

Firstly, we associate to every vertex in  $G_n$  a location on the negative half-line such that the youngest vertex (i.e., vertex  $n$ ) is located at the origin and the distance between vertex  $j-1$  and vertex  $j$  is given by  $1/j$ . In particular, the vertex labelled  $v$  is located at  $s_n(v) := -\sum_{j=v}^{n-1} \frac{1}{j}$ , the location of the oldest vertex scales like  $-\log n$ , and vertices with label at most  $\lfloor \epsilon n \rfloor$ , which we remove when damaging the network, are asymptotically located to the left of  $\log \epsilon$ . The location of a vertex is determined by its age in the network with old vertices being located further left than young vertices. As the graph size increases, the location of any fixed vertex moves to the left and the vertex locations  $(s_n(v) : v \in \{1, \dots, n\})$  become dense everywhere on the negative half-line.

We run an exploration from vertex  $v_0 \in G_n^\epsilon$  and successively create particles in the branching random walk that approximate the discovered vertices. We stop as soon

as there is no longer a one-to-one correspondence between the nodes in the two processes. For example, this could happen if in the network a vertex is rediscovered and the explored subgraph is no longer a tree. A careful analysis, carried out in Section 4.1 below, shows that when the order in which vertices are explored is chosen in a suitable way, then we do not stop until either the whole component is discovered or at least  $c_n$  vertices have been found, where  $\lim_{n \rightarrow \infty} c_n^2/n = 0$ .

To start, we place a particle at the location of vertex  $v_0$  and declare it to be the root of the branching random walk. Then we explore all direct neighbours of  $v_0$  in  $G_n^\epsilon$ . The locations of the particles in the first generation of the branching random walk are chosen to approximate the locations of these direct neighbours. To this end, we distinguish offspring located to the left and right of  $v_0$ . For a given interval  $[a, b]$  on the left of  $s_n(v_0)$ , i.e.  $[a, b] \subseteq [\log \epsilon, s_n(v_0)]$ , the number of vertices located in  $[a, b]$  is a sum of independent Bernoulli random variables by the definition of the model. The probability that  $v_0$  has a direct neighbour labelled  $u < v_0$ , is given by

$$\mathbb{P}(\Delta \mathcal{Z}[u, v_0 - 1] = 1) = \frac{1}{v_0 - 1} \mathbb{E}[f(\mathcal{Z}[u, v_0 - 1])] \approx \int_{s_n(u-1) - s_n(v_0)}^{s_n(u) - s_n(v_0)} \beta e^{(1-\gamma)t} dt, \quad (1.8)$$

where  $\mathcal{Z}[u, v]$  is the indegree of vertex  $u$  at time  $v$ ,  $\Delta \mathcal{Z}[u, v_0 - 1] = \mathcal{Z}[u, v_0] - \mathcal{Z}[u, v_0 - 1]$ , and the approximation follows from the fact that the process  $(f(\mathcal{Z}[u, n]) \prod_{j=u}^{n-1} \frac{1}{1+\gamma/j} : n \geq u)$  is a martingale; see Lemma 3.8 below. Since the location of  $u$  can be written as  $s_n(v_0)$  plus the displacement  $s_n(u) - s_n(v_0)$ , asymptotically, we can approximate the displacements of the direct neighbours of  $v_0$  on its left by the points of a Poisson point process  $\Pi$  with intensity measure  $\beta e^{(1-\gamma)t}$  on  $(-\infty, 0]$  that lie in  $[\log \epsilon - s_n(v_0), 0]$ . We emphasise that  $\Pi$  describes the displacements, not the particle locations. Hence, in the branching random walk, the relative positions of the offspring to the left of a particle with location  $\lambda$  are given by the points of  $\Pi$  that lie in  $[\log \epsilon - \lambda, 0]$ .

In the next step, we motivate the point process that describes the relative positions of the offspring on the right in the branching random walk. Note that in the network every direct neighbour  $u$  of  $v_0$  with  $u > v_0$  increases the indegree of  $v_0$  and therefore the probability that  $v_0$  has further offspring on its right. The distance between the  $i$ -th and  $(i+1)$ -st right neighbour of  $v_0$  in the network is given by

$$T_{v_0}[i] := \sup \left\{ \sum_{j=k}^l \frac{1}{j} : \mathcal{Z}[v_0, k] = i = \mathcal{Z}[v_0, l] \right\}.$$

Suppose the  $i$ -th neighbour of  $v_0$  is born at time  $k$ . For given  $t > 0$ , we have

$$\mathbb{P}(T_{v_0}[i] > t \mid \mathcal{Z}[v_0, k-1] < \mathcal{Z}[v_0, k] = i) = \mathbb{P}(\mathcal{Z}[v_0, l] = i \mid \mathcal{Z}[v_0, k] = i), \quad (1.9)$$

where  $l$  is the smallest integer with  $\sum_{j=k}^l \frac{1}{j} > t$ . Plugging in the connection probabilities, we deduce that (1.9) is equal to

$$\prod_{j=k}^{l-1} \left( 1 - \frac{f(i)}{j} \right) \approx \exp \left( -f(i) \sum_{j=k}^{l-1} \frac{1}{j} \right) \approx \exp(-f(i)t).$$

Hence, the distance between the  $i$ -th and  $(i+1)$ -st right neighbour of  $v_0$  is approximately exponentially distributed with rate  $f(i)$ . For a precise statement, see Lemma 3.2 below. Consequently, the displacements of the direct neighbours on the right of  $v_0$  are well approximated by the jump times in  $[0, -s_n(v_0)]$  of the pure jump Markov process  $Z = (Z_t : t \geq 0)$  that starts in  $Z_0 = 0$  and jumps from  $i$  to  $i+1$  after an exponential waiting time with rate  $f(i)$ , independently of the previous jumps. Therefore, in the branching

random walk, the relative positions of the offspring to the right of the root located in  $\lambda$  are given by jump times of  $Z$  in  $[0, -\lambda]$ .

When the exploration is continued, the information gathered from the already explored neighbourhood leads to a size-biasing effect. Indeed, in the network the edges between a vertex  $v$  and its direct neighbours on the right  $u > v$  are not independent. If  $v$  was discovered as a direct neighbour of a vertex  $w$  on its right, i.e.  $w > v$ , then we already know that  $v$  has indegree at least one. Consequently, we expect  $v$  to have more direct neighbours on its right than without this information. Mathematically, this leads to a size biasing effect and the displacements of particles on the right of  $v$  are given by the jump times in  $[0, -s_n(v)]$  of  $Z$  started in *one* instead of zero. In contrast, if  $v$  was discovered as a direct neighbour of a vertex  $w$  with smaller label, i.e.  $w < v$ , then we do not have that information and the displacements are again the jump times in  $[0, -s_n(v)]$  of  $Z$  started in zero. Similarly, for the direct neighbours on the left of  $v$ , there is no size-biasing effect as a consequence of the independence between the edges on the left. Of course, there are several further dependencies coming from the previously explored subgraph. However, we show in Section 4 that the error accrued by adjusting only for the immediate parent is asymptotically negligible when we discover not more than  $c_n$  vertices, where  $\lim_{n \rightarrow \infty} c_n^3/n = 0$ .

To be able to use different offspring distributions depending on the relative location of the parent, each vertex is equipped with a mark  $\alpha$  in  $\{\ell, r\}$  to indicate the relative location of the parent, where the non-numerical symbols  $\ell$  and  $r$  stand for ‘left’ and ‘right’, respectively. The relative positions of the offspring can be generated as the points of  $\Pi$  on  $(-\infty, 0]$  and the jump times of  $Z$  (with initial state depending on the mark) on  $[0, \infty)$ . All offspring particles located on the left of  $\log \epsilon$  or on the right of 0 are immediately removed. In other words, the approximating tree is the genealogical tree of a two-type branching random walk with two killing boundaries.

An equivalent description is as a multitype branching process with type space  $\Phi := [\log \epsilon, 0] \times \{\ell, r\}$ , where the first component indicates the location of a particle and the second indicates its mark. Whilst the branching random walk interpretation offers more intuition, the two killing boundaries make the mathematical analysis difficult. Hence in our analysis, we will use the interpretation of the process as multitype branching process with the larger typespace  $\Phi$ .

## 1.7 Main ideas of the proofs

Understanding the local neighbourhood of vertices in the network is the key to many of its properties. As in [15] the survival probability of the approximating killed branching random walk is equal to the asymptotic relative size  $\zeta$  of the largest component. This result allows us to determine, for example, the critical parameter for percolation from knowledge when the percolated branching process has a positive survival probability. To form the percolated branching process with retention probability  $p$  from the original process every particle is kept with probability  $p$  and removed together with its line of descent with probability  $1 - p$  independently of all other particles.

It is instructive to continue the comparison of the damaged and undamaged networks in the setup of this branching process. In [15], where the undamaged network is analysed, the branching random walk has only *one* killing boundary on the right. It turns out that on the set of survival the leftmost particle drifts away from the killing boundary, such that it does not feel the boundary anymore. As a consequence, the un-killed process carries all information needed to determine whether the killed branching random walk survives with positive probability and, therefore, whether the network has a giant component. The *two* killing boundaries in the branching random walk describing the damaged network prevent us from using this analogy; every particle is exposed

to the threat of absorption.

To survive indefinitely, a genealogical line of descent has to move within the (space-time) strip  $[\log \epsilon, 0] \times \mathbb{N}_0$ . To understand the optimal strategy for survival observe that, in the network with strong preferential attachment, old vertices typically have a large degree and therefore are connected to many young vertices, while young vertices themselves have only a few connections. This means that in the branching random walk without killing, particles produce many offspring to the right, but only a few to the left. Hence, if a particle is located near the left killing boundary it represents an old vertex in the graph and is very fertile, but its offspring are mostly located further to the right and are therefore less fertile. A particle near the right killing boundary, however, represents a young vertex and has itself a small number of offspring, which then however have a good chance of being fertile since they are necessarily located further left in the strip. As a result, the optimal survival strategy for a particle is to have an ancestral line of particles whose locations are alternating between positions near the left and the right killing boundary. This intuition is the basis for our proofs.

Continuing more formally, for the proof of Theorem 1.1 we show that positivity of the survival probability can be characterised in terms of the largest eigenvalue  $\rho_\epsilon$  of an operator that describes the spatial distribution of offspring of a given particle. More precisely, the branching random walk survives percolation with retention parameter  $p$  if its growth rate  $p\rho_\epsilon$  exceeds the value one, so that  $p_c(\epsilon) = 1/\rho_\epsilon$ . Our intuition allows us to guess the form of the corresponding eigenfunction, which, relatively to the particle density, has its mass concentrated in two bumps near the left and right killing boundary. From this guess we obtain sufficiently accurate estimates for the largest eigenvalue, and therefore for the critical percolation parameter, as long as the preferential attachment effect is strong enough. This is the case if  $\gamma \geq \frac{1}{2}$ , allowing us to prove Theorem 1.1.

By contrast, for  $\gamma < \frac{1}{2}$  we know that the network is not robust, i.e. we have  $p_c(0) > 0$ . It would be of interest to understand the behaviour of  $p_c(\epsilon) - p_c(0)$  as  $\epsilon \downarrow 0$ . Our methods can be applied to this case, but the resulting bounds are very rough. The reason is that in this regime the preferential attachment is much weaker, and the intuitive idea underlying our estimates gives a less accurate picture.

The idea for the proof of Theorem 1.3 is based on the branching process comparison, too. To bound the probability that two typical vertices  $V$  and  $W$  are connected by a path of length at most  $h$ , we look at the expected number of such paths. That is given by the number of vertices at distance at most  $h - 1$  from  $V$  multiplied by the probability that such a vertex connects to  $W$ . By our branching process heuristics, the number of vertices at distance at most  $h - 1$  from  $V$  can be approximated by the number of particles in the first  $h - 1$  generations of the branching random walk, which is of order  $\rho_\epsilon^h$  where  $\rho_\epsilon = 1/p_c(\epsilon)$  as before. The probability of connecting any vertex with label at least  $\epsilon n$  to  $W$  is bounded from above by  $f(m)/\epsilon n$ , where  $m$  is the maximal degree in the network. Since  $m = o(n)$  by Theorem 1.2, this implies that the probability of a connection between  $V$  and  $W$  is bounded from above by  $\exp(h \log(1/p_c(\epsilon))) - \log n + o(\log n)$  and therefore goes to zero if  $h \leq (1 - \delta) \log n / \log(1/p_c(\epsilon))$ ,  $\delta > 0$ , which yields the result.

Theorem 1.2 is relatively soft by comparison. The independence of the indegrees of distinct vertices allows us to study them separately and we again use the continuous approximation to describe the expected empirical indegree evolution. The limit theorem for the empirical distribution itself follows from a standard concentration argument. The asymptotic result for the maximal degrees is only slightly more involved and is based on fairly standard extreme value arguments.

## 1.8 Overview

The outline of this article is as follows. We start with the main steps of the proofs in Section 2. The multitype branching process which locally approximates a connected component in the network is defined in Section 2.1 and its key properties are stated. The main part of the proof of Theorem 1.1 then follows in Section 2.2. The analysis of the multitype branching process is conducted in Section 2.3. We study an operator associated with this process in Section 2.3.1 and derive necessary and sufficient conditions for its survival in Section 2.3.2. Sections 3.1 and 3.2 are devoted to the study of the topology of the damaged graph. In Section 3.1 the typical and maximal degree of vertices is analysed. In Section 3.2 typical distances are studied. The couplings between the network and the approximating branching process that underlie our proofs are provided in Section 4. We then look at model variations in Section 5. The derivation of Theorem 1.4 from Theorem 1.1 is presented in Section 5.1. This is the only section which requires consideration of non-linear attachment rules. We finish in Section 5.2 by studying the question of vulnerability in other network models.

## 2 Connectivity and branching processes

In this section, we restrict our attention to linear attachment rules  $f(k) = \gamma k + \beta$ , for  $\gamma \in [0, 1)$  and  $\beta > 0$ , and let  $\epsilon$  be a fixed value in  $(0, 1)$ . The goal of this section is to prove Theorem 1.1. To this end, we couple the local neighbourhood of a vertex in  $G_n^\epsilon$  to a multitype branching process. The branching process is introduced in Section 2.1 and Theorem 1.1 is deduced in Section 2.2. Properties of the branching processes which are needed in the analysis are proved in Section 2.3. The proof of the coupling between network and branching process is deferred to Section 4.

### 2.1 The approximating branching process

As explained in Section 1.6 the local neighbourhood of a vertex in  $G_n^\epsilon$  can be approximated by a multitype branching process with type space  $\Phi = [\log \epsilon, 0] \times \{\ell, r\}$ . A typical element of  $\Phi$  is denoted by  $\phi = (\lambda, \alpha)$ . The intuitive picture is that  $\lambda$  encodes the spatial position of the particle which we call *location*. The second coordinate  $\alpha$  indicates on which side of the particle its parent is located and we refer to  $\alpha$  as the *mark*. In view of (1.8), a particle of type  $(\lambda, \alpha) \in \Phi$  produces offspring to its left with displacements having the same distribution as those points of the Poisson point process  $\Pi$  on  $(-\infty, 0]$  with intensity measure

$$\beta e^{(1-\gamma)t} \mathbb{1}_{(-\infty, 0]}(t) dt \quad (2.1)$$

that lie in  $[\log \epsilon - \lambda, 0]$ . Since these offspring have their parent on the right, they are of mark  $r$ .

Recall from Section 1.6 that we denote by  $Z$  an increasing, integer-valued process, which jumps from  $i$  to  $i + 1$  after an exponential waiting time with rate  $f(i)$ , independently of the previous jumps. We write  $P$  for the distribution of  $Z$  started in zero and  $E$  for the corresponding expectation. By  $(\hat{Z}_t: t \geq 0)$  we denote a version of the process started in  $\hat{Z}_0 = 1$  under the measure  $P$ .

The distribution of the offspring to the right depends on the mark of the parent. As motivated in Section 1.6, when the particle is of type  $(\lambda, \ell)$ , then the displacements of the offspring follow the same distribution as the jump times of  $(Z_t: t \in [0, -\lambda])$ , but when the particle is of type  $(\lambda, r)$ , then the displacements follow the same distribution as the jump times of  $(\hat{Z}_t: t \in [0, -\lambda])$ . All offspring on the right have their parent on the left, so their mark is  $\ell$ . Observe that the chosen offspring distributions ensure that new particles have again a location in  $[\log \epsilon, 0]$ . The offspring distribution to the right is *not*



a Poisson point process: The more particles are born, the higher the rate at which new particles arrive.

We call the branching process thus constructed the *idealized branching process* (IBP). It can be interpreted as a labelled tree, where every node represents a particle and is connected to its children and (apart from the root) to its parent. We equip node  $x$  with label  $\phi(x) = (\lambda(x), \alpha(x))$ , where  $\lambda(x)$  denotes its location and  $\alpha(x)$  its mark, and write  $|x|$  for the generation of  $x$ . To obtain a branching process approximation to  $G_n^\epsilon(p)$ , we define the *percolated* IBP by associating to every offspring in the IBP an independent Bernoulli( $p$ ) random variable. If the random variable is zero, we delete the offspring together with its line of descent. If it equals one, then the offspring is retained in the percolated IBP.

Let  $S^\epsilon$  be a random variable with

$$\mathbb{P}(-S^\epsilon \leq t) = \frac{1}{1-\epsilon}(1 - e^{-t}), \quad \text{for } t \in [0, -\log \epsilon]. \quad (2.2)$$

Denote by  $\zeta^\epsilon(p)$  the survival probability of the tree which is with probability  $p$  equal to the percolated IBP started with one particle of mark  $\ell$  and location  $S^\epsilon$  and equals the empty tree otherwise. Let  $C_n^\epsilon(p)$  be a connected component in  $G_n^\epsilon(p)$  of maximal size.

**Theorem 2.1.** *For all  $\epsilon \in (0, 1)$  and  $p \in (0, 1]$ , in probability,*

$$\frac{|C_n^\epsilon(p)|}{\mathbb{E}|V_n^\epsilon(p)|} \rightarrow \zeta^\epsilon(p)/p \quad \text{as } n \rightarrow \infty.$$

The proof of Theorem 2.1 is postponed to Section 4. The theorem describes the asymptotic size of the largest component in the network in terms of the survival probability of the percolated IBP. To make use of this connection, we have to understand the branching process.

For any measurable, complex-valued, bounded function  $g$  on  $\Phi$ , and  $\phi \in \Phi$ , let

$$A_p g(\phi) := E_{\phi,p} \left[ \sum_{|x|=1} g(\lambda(x), \alpha(x)) \right],$$

where the expectation  $E_{\phi,p}$  refers to the percolated IBP starting with a single particle of type  $\phi$ , percolated with retention parameter  $p$ . We write  $A = A_1$  for the operator corresponding to the unpercolated branching process and  $E_\phi := E_{\phi,1}$ . Recall that all quantities associated with the IBP, and in particular  $A_p$ , depend on the fixed value of  $\epsilon$ . We denote by  $C(\Phi)$  the complex Banach space of continuous functions on  $\Phi$  equipped with the supremum norm. The following proposition, which summarizes properties of  $A_p$ , is proved in Section 2.3.1.

**Proposition 2.2.** *For all  $\epsilon \in (0, 1)$  and  $p \in (0, 1]$ , the operator  $A_p: C(\Phi) \rightarrow C(\Phi)$  is linear, strictly positive and compact with spectral radius  $\rho_\epsilon(A_p) \in (0, \infty)$ . Moreover,  $A_p = pA$  and  $\rho_\epsilon(A_p) = p\rho_\epsilon(A)$ .*

The survival probability of the percolated IBP has the following property.

**Theorem 2.3.** *For all  $\epsilon \in (0, 1)$  and  $p \in (0, 1]$*

$$\zeta^\epsilon(p) > 0 \quad \Leftrightarrow \quad \rho_\epsilon(A_p) > 1.$$

Theorem 2.3 is proved in Section 2.3.2. Combined with Theorem 2.1 and Proposition 2.2 it gives a characterisation of the critical percolation parameter for the network  $(G_n^\epsilon(p): n \in \mathbb{N})$ .

**Corollary 2.4.** *The network  $(G_n^\epsilon(p): n \in \mathbb{N})$  has a giant component if and only if the retention probability  $p$  satisfies  $p > \rho_\epsilon(A)^{-1}$ .*

Notice that the corollary implies that  $(G_n^\epsilon: n \in \mathbb{N})$  has no giant component when  $\rho_\epsilon(A) \leq 1$ . Moreover, the first statement of Theorem 1.1 follows from the corollary by taking  $p_c(\epsilon) = \rho_\epsilon(A)^{-1} \wedge 1$ .

To complete the proof of Theorem 1.1, it remains to estimate the spectral radius  $\rho_\epsilon(A)$ . This estimation is performed in Section 2.2 below using that (see, e.g., Theorem 45.1 in [22]) for a linear and bounded operator  $A$  on a complex Banach space, the spectral radius is given by

$$\rho(A) = \lim_{n \rightarrow \infty} \|A^n\|^{\frac{1}{n}} = \inf\{\|A^n\|^{\frac{1}{n}} : n \in \mathbb{N}\}. \quad (2.3)$$

By the definition of the Poisson point process  $\Pi$  in (2.1), the intensity measure of  $\Pi$  equals

$$\mathbb{1}_{(-\infty, 0]}(t) M(dt), \quad \text{for } M(dt) := \beta e^{(1-\gamma)t} dt.$$

We denote by  $\Pi^\ell$  the point process given by the jump times of  $(Z_t: t \geq 0)$  and by  $\Pi^r$  the point process given by the jump times of  $(\tilde{Z}_t: t \geq 0)$ . A simple computation (cf. Lemma 1.12 in [15]) shows that with  $M^\alpha(dt) := a_\alpha e^{\gamma t} dt$ , where  $a_\ell = \beta$  and  $a_r = \gamma + \beta$ , the intensity measure of  $\Pi^\alpha$  is given by  $\mathbb{1}_{[0, \infty)}(t) M^\alpha(dt)$  for  $\alpha \in \{\ell, r\}$ . Hence, for any bounded, measurable function  $g$  on  $\Phi$  and  $(\lambda, \alpha) \in \Phi$ ,

$$\begin{aligned} Ag(\lambda, \alpha) &= E_{(\lambda, \alpha)} \left[ \sum_{|x|=1} g(\lambda(x), \alpha(x)) \right] \\ &= \int_{\log \epsilon - \lambda}^0 g(\lambda + t, r) \beta e^{(1-\gamma)t} dt + \int_0^{-\lambda} g(\lambda + t, \ell) a_\alpha e^{\gamma t} dt. \end{aligned} \quad (2.4)$$

## 2.2 Proof of Theorem 1.1

Subject to the considerations of the previous section, Theorem 1.1 follows from the following proposition.

**Proposition 2.5.** (a) *If  $\gamma = \frac{1}{2}$ , then  $\frac{1}{\gamma + \beta} \frac{1}{\log(1/\epsilon)} \leq \rho_\epsilon(A)^{-1} \leq \frac{1}{\beta} \frac{1}{\log(1/\epsilon)}$ .*

(b) *If  $\gamma > \frac{1}{2}$ , then*

$$\begin{aligned} &\left(1 + \log(\epsilon^{1-2\gamma}) \epsilon^{\gamma-1/2} + [\log(\epsilon^{1-2\gamma}) \epsilon^{\gamma-1/2}]^2\right)^{-1/2} \\ &\leq \frac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1} \epsilon^{-\gamma+1/2} \rho_\epsilon(A)^{-1} \leq (1 - \epsilon^{\gamma-1/2})^{-1}. \end{aligned}$$

*Proof of Proposition 2.5 (a).* For  $h_0 \in C([\log \epsilon, 0])$ , the complex Banach space of continuous functions on  $[\log \epsilon, 0]$ , let

$$\bar{A}h_0(\lambda) := \int_{\log \epsilon - \lambda}^0 h_0(\lambda + t) e^{t/2} dt + \int_0^{-\lambda} h_0(\lambda + t) e^{t/2} dt \quad \text{for } \lambda \in [\log \epsilon, 0].$$

Note that  $A$  and  $\bar{A}$  map real-valued functions to real-valued functions and nonnegative functions to nonnegative functions, and they are monotone. For  $A$  this observation implies

$$\|A^n\| = \sup\{\|A^n g\| : g \in C(\Phi), g(\phi) \in [0, 1] \text{ for all } \phi \in \Phi\} = \|A^n \mathbf{1}\|, \quad (2.5)$$

where  $\|A^n g\| = \sup\{|A^n g(\phi)| : \phi \in \Phi\}$  and  $\mathbf{1}$  denotes the constant function with value 1. Combining (2.3) and (2.5), we deduce that  $\rho_\epsilon(A) = \lim_{n \rightarrow \infty} \|A^n \mathbf{1}\|^{\frac{1}{n}}$ . The same argument shows that

$$\rho_\epsilon(\bar{A}) = \lim_{n \rightarrow \infty} \|\bar{A}^n \mathbf{1}\|^{\frac{1}{n}}. \quad (2.6)$$

Defining  $h(\lambda, \alpha) := h_0(\lambda)$  for all  $(\lambda, \alpha) \in \Phi$ , (2.4) yields for  $h_0 \in C([\log \epsilon, 0])$  with  $h_0 \geq 0$ ,

$$\beta \bar{A} h_0(\lambda) \leq A h(\lambda, \alpha) \leq (\gamma + \beta) \bar{A} h_0(\lambda) \quad \text{for all } (\lambda, \alpha) \in \Phi.$$

In particular, by the monotonicity and linearity of  $A$  and  $\bar{A}$ ,

$$\beta^n \bar{A}^n \mathbf{1}(\lambda) \leq A^n \mathbf{1}(\lambda, \alpha) \leq (\gamma + \beta)^n \bar{A}^n \mathbf{1}(\lambda) \quad \text{for } (\lambda, \alpha) \in \Phi, n \in \mathbb{N},$$

implying  $\rho_\epsilon(A) \in [\beta \rho_\epsilon(\bar{A}), (\gamma + \beta) \rho_\epsilon(\bar{A})]$ . To complete the proof it suffices to show that  $\rho_\epsilon(\bar{A}) = \log(1/\epsilon)$ , which we can achieve by ‘guessing’ the principal eigenfunction of  $\bar{A}$ . Indeed, the result follows from (2.6) and

$$\bar{A}^{n+1} \mathbf{1}(\lambda) = 2(1 - \epsilon^{1/2})(\log(1/\epsilon))^n e^{-\lambda/2} \quad \text{for } \lambda \in [\log \epsilon, 0], n \in \mathbb{N}_0.$$

We show the latter identity by induction over  $n$ . For  $n = 0$ ,

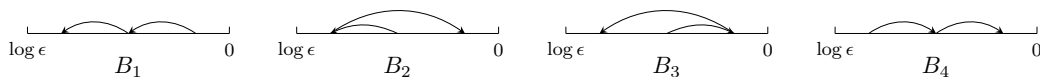
$$\bar{A} \mathbf{1}(\lambda) = \int_{\log \epsilon - \lambda}^0 e^{t/2} dt + \int_0^{-\lambda} e^{t/2} dt = 2(1 - e^{-\lambda/2} \epsilon^{1/2} + e^{-\lambda/2} - 1) = 2(1 - \epsilon^{1/2}) e^{-\lambda/2}.$$

Moreover, with  $h_0(\lambda) := e^{-\lambda/2}$  we have

$$\bar{A} h_0(\lambda) = \int_{\log \epsilon - \lambda}^0 e^{-(\lambda+t)/2} e^{t/2} dt + \int_0^{-\lambda} e^{-(\lambda+t)/2} e^{t/2} dt = e^{-\lambda/2} \log(1/\epsilon).$$

Thus,  $\lambda \mapsto e^{-\lambda/2}$  is an eigenfunction of  $\bar{A}$  with eigenvalue  $\log(1/\epsilon)$  and the proof is complete.  $\square$

*Proof of lower bound in Proposition 2.5 (b).* We analyse the ancestral lines of particles in the branching process at a fixed time  $n \geq 2$ . Going back two steps in the ancestral line of every particle alive we can divide the population at time  $n$  in four groups depending on the relative positions of parent and child in the transitions from generation  $n - 2$  to  $n - 1$  and from  $n - 1$  to  $n$ : (1) in both steps the child is to the left of its parent, (2) in the first step the child is to the left and in the second it is to the right of its parent, (3) first right, then left, (4) in both steps the child is to the right of its parent. The cases are depicted in Figure 3.



**Figure 3:** Possible genealogy of a particle contributing to the respective operators.

We denote by  $B_i$ ,  $i \in \{1, \dots, 4\}$ , the operators corresponding to the four scenarios. Using the point processes  $\Pi$ ,  $\Pi^\ell$  and  $\Pi^r$  this means, for any bounded, measurable

function  $g$  on  $\Phi$  and  $(\lambda, \alpha) \in \Phi$ ,

$$\begin{aligned} B_1 g(\lambda, \alpha) &:= \mathbb{E} \left[ \sum_{\substack{\mathbf{p} \in \Pi \\ \log \epsilon - \lambda \leq \mathbf{p}}} \sum_{\substack{\mathbf{q} \in \Pi \\ \log \epsilon - \lambda - \mathbf{p} \leq \mathbf{q}}} g(\lambda + \mathbf{p} + \mathbf{q}, r) \right], \\ B_2 g(\lambda, \alpha) &:= \mathbb{E} \left[ \sum_{\substack{\mathbf{p} \in \Pi \\ \log \epsilon - \lambda \leq \mathbf{p}}} \sum_{\substack{\mathbf{q} \in \Pi^r \\ \mathbf{q} \leq -(\lambda + \mathbf{p})}} g(\lambda + \mathbf{p} + \mathbf{q}, \ell) \right], \\ B_3 g(\lambda, \alpha) &:= \mathbb{E} \left[ \sum_{\substack{\mathbf{p} \in \Pi^\alpha \\ \mathbf{p} \leq -\lambda}} \sum_{\substack{\mathbf{q} \in \Pi \\ \log \epsilon - \lambda - \mathbf{p} \leq \mathbf{q}}} g(\lambda + \mathbf{p} + \mathbf{q}, r) \right], \\ B_4 g(\lambda, \alpha) &:= \mathbb{E} \left[ \sum_{\substack{\mathbf{p} \in \Pi^\alpha \\ \mathbf{p} \leq -\lambda}} \sum_{\substack{\mathbf{q} \in \Pi^\ell \\ \mathbf{q} \leq -(\lambda + \mathbf{p})}} g(\lambda + \mathbf{p} + \mathbf{q}, \ell) \right]. \end{aligned}$$

Intuitively, going back the ancestral line of a typical particle in the population at a late time, for a few generations the ancestral particles may be in group (4), because of the high fertility of particles positioned near the left boundary of  $[\log \epsilon, 0]$ . But this behaviour is not sustainable, as after a few generations in this group the offspring particle will typically be near the right end of the interval and will therefore be pushed into the right killing boundary so that it is likely to die out. Over a longer period the ancestral particles are much more likely to be in groups (2) and (3), as this behaviour is sustainable over long periods when the ancestral line is hopping more or less regularly between positions near the left and the right boundary of the interval  $[\log \epsilon, 0]$ . A similar pattern can also be observed when studying typical paths in the random graph model; see our discussion in Section 1.7. The aim is now to turn this heuristics into useful bounds on high iterates of the operator  $A$ .

It is useful to understand how the operators  $B_i$  act on the constant function  $\mathbf{1}$  as well as on the functions  $g_1(\lambda, \alpha) := e^{-\gamma\lambda}$  and  $g_2(\lambda, \alpha) := e^{-(1-\gamma)\lambda}$ . We can write

$$B_3 g(\lambda, \alpha) = \int_0^{-\lambda} \int_{\log \epsilon - \lambda - t}^0 g(\lambda + t + s, r) M(ds) M^\alpha(dt),$$

where  $M(dt) = \beta e^{(1-\gamma)t} dt$  and  $M^\alpha(dt) = a_\alpha e^{\gamma t} dt$  with  $a_\alpha \leq \gamma + \beta$  are the intensity measures of the point processes  $\Pi$  and  $\Pi^\alpha$ . From this we obtain, for  $(\lambda, \alpha) \in \Phi$ ,

$$\begin{aligned} B_3 \mathbf{1}(\lambda, \alpha) &\leq \int_{-\infty}^{-\lambda} \int_{-\infty}^0 M(ds) M^r(dt) = \frac{\beta(\gamma + \beta)}{\gamma(1 - \gamma)} e^{-\gamma\lambda}, \\ B_3 g_1(\lambda, \alpha) &\leq \int_{-\infty}^{-\lambda} \int_{\log \epsilon - t - \lambda}^{\infty} e^{-\gamma(\lambda + t + s)} M(ds) M^r(dt) = \frac{\beta(\gamma + \beta)}{(2\gamma - 1)^2} \epsilon^{1-2\gamma} e^{-\gamma\lambda}, \\ B_3 g_2(\lambda, \alpha) &\leq \int_{-\infty}^{-\lambda} \int_{\log \epsilon}^0 e^{-(1-\gamma)(\lambda + t + s)} M(ds) M^r(dt) = \frac{\beta(\gamma + \beta)}{2\gamma - 1} \log(1/\epsilon) e^{-\gamma\lambda}. \end{aligned}$$

Moreover, similarly elementary calculations for  $B_1$ ,  $B_2$  and  $B_4$  imply

$$\begin{aligned} B_1 \mathbf{1}(\lambda, \alpha) &\leq \frac{\beta^2}{(1-\gamma)^2}, & B_2 \mathbf{1}(\lambda, \alpha) &\leq \frac{\beta(\gamma + \beta)}{\gamma(2\gamma - 1)} \epsilon^{1-2\gamma} e^{-(1-\gamma)\lambda}, \\ B_4 \mathbf{1}(\lambda, \alpha) &\leq \frac{\beta(\gamma + \beta)}{\gamma} \log(1/\epsilon) e^{-\gamma\lambda}; \end{aligned}$$

and

$$\begin{aligned} B_1 g_1(\lambda, \alpha) &\leq \frac{\beta^2}{2\gamma - 1} \log(1/\epsilon) \epsilon^{1-2\gamma} e^{-(1-\gamma)\lambda}, & B_1 g_2(\lambda, \alpha) &\leq \beta^2 (\log \epsilon)^2 e^{-(1-\gamma)\lambda}, \\ B_2 g_1(\lambda, \alpha) &\leq \frac{\beta(\gamma + \beta)}{2\gamma - 1} \log(1/\epsilon) \epsilon^{1-2\gamma} e^{-(1-\gamma)\lambda}, & B_2 g_2(\lambda, \alpha) &\leq \frac{\beta(\gamma + \beta)}{(2\gamma - 1)^2} \epsilon^{1-2\gamma} e^{-(1-\gamma)\lambda}, \\ B_4 g_1(\lambda, \alpha) &\leq \beta(\gamma + \beta) (\log \epsilon)^2 e^{-\gamma\lambda}, & B_4 g_2(\lambda, \alpha) &\leq \frac{\beta(\gamma + \beta)}{2\gamma - 1} \log(1/\epsilon) e^{-\gamma\lambda}. \end{aligned}$$

Summarising, there exists  $C_\epsilon > 0$  such that  $B_i \mathbf{1}(\phi) \leq C_\epsilon g_1(\phi)$  for all  $i \in \{1, \dots, 4\}$ ,  $\phi \in \Phi$ , and denoting

$$b_{\text{sm}} := b_1 := b_4 := \beta(\gamma + \beta)(\log \epsilon)^2, \quad b_{\text{bg}} := b_2 := b_3 := \frac{\beta(\gamma + \beta)}{(2\gamma - 1)^2} \epsilon^{1-2\gamma},$$

where bg stands for ‘big’ and sm for ‘small’, we have

$$\begin{aligned} B_i g_1(\phi) &\leq b_{\text{bg}} \log(\epsilon^{1-2\gamma}) g_2(\phi), & B_i g_2(\phi) &\leq b_i g_2(\phi) && \text{for } i \in \{1, 2\}, \\ B_i g_1(\phi) &\leq b_i g_1(\phi), & B_i g_2(\phi) &\leq b_{\text{bg}} \log(\epsilon^{1-2\gamma}) \epsilon^{2\gamma-1} g_1(\phi) && \text{for } i \in \{3, 4\}. \end{aligned}$$

Using that by definition  $A^2 = \sum_{i=1}^4 B_i$ , our estimate for  $B_i \mathbf{1}$  yields

$$A^{2(n+1)} \mathbf{1}(\phi) = \sum_{i_0, \dots, i_n \in \{1, \dots, 4\}} B_{i_n} \circ \dots \circ B_{i_0} \mathbf{1}(\phi) \leq 4C_\epsilon \sum_{i_1, \dots, i_n \in \{1, \dots, 4\}} B_{i_n} \circ \dots \circ B_{i_1} g_1(\phi). \quad (2.7)$$

Up to constants, the estimates for  $B_3$  and  $B_4$  preserve  $g_1$  but change  $g_2$  into  $g_1$ , whereas the estimates for  $B_1$  and  $B_2$  preserve  $g_2$  and change  $g_1$  into  $g_2$ . Hence, we split the sequence of indices into blocks containing only 1 or 2 and blocks containing only 3 or 4. We write  $m$  for the number of blocks,  $k_j$  for the length of block  $j$  and  $\bar{k}_j := \sum_{i=1}^{j-1} k_i + 1$  for the first index in block  $j$ . Then

$$\sum_{i_1, \dots, i_n \in \{1, \dots, 4\}} B_{i_n} \circ \dots \circ B_{i_1} g_1(\phi) = \sum_{m=1}^{n+1} \sum_{\substack{k_1 + \dots + k_m = n \\ k_1 \in \mathbb{N}_0, k_2, \dots, k_m \in \mathbb{N}}} \sum_{(i_1, \dots, i_n)} B_{i_n} \circ \dots \circ B_{i_1} g_1(\phi),$$

where the last sum is over all sequences of indices  $(i_1, \dots, i_n)$  with  $i_{\bar{k}_j}, \dots, i_{\bar{k}_{j+1}-1} \in \{3, 4\}$  for  $j$  odd and  $i_{\bar{k}_j}, \dots, i_{\bar{k}_{j+1}-1} \in \{1, 2\}$  for  $j$  even. We insist that formally the first block contains the indices 3 or 4 — the case that this does not hold is covered by  $k_1 = 0$ . Hence, in the first block, operators  $B_3$  and  $B_4$  encounter  $g_1$ , which is preserved. To determine the constants, we only have to keep track of how often  $B_4$  is used; we call this number  $l_1$ . The first operator belonging to a new block  $j$  causes a factor  $b_{\text{bg}} \log(\epsilon^{1-2\gamma})$  and if the change is from a  $\{1, 2\}$  to a  $\{3, 4\}$  block, then an additional  $\epsilon^{2\gamma-1}$  is obtained. For the subsequent steps within block  $j$ , we again have to track how often the operator causing the smaller constant  $b_{\text{sm}}$ ,  $B_1$  or  $B_4$ , is used. This number is called  $l_j$ . After applying all  $n$  operators, the function  $g_1(\phi) \mathbb{1}_{\text{odd}}(m) + g_2(\phi) \mathbb{1}_{\text{even}}(m)$  remains and we bound it by  $\epsilon^{-\gamma}$ . This procedure yields

$$\begin{aligned} &\sum_{i_1, \dots, i_n \in \{1, \dots, 4\}} B_{i_n} \circ \dots \circ B_{i_1} g_1(\phi) \\ &\leq \sum_{m=1}^{n+1} \sum_{\substack{k_1 + \dots + k_m = n \\ k_1 \in \mathbb{N}_0, k_2, \dots, k_m \in \mathbb{N}}} b_{\text{bg}}^{m-1} (\log(\epsilon^{1-2\gamma}))^{m-1} \epsilon^{(2\gamma-1)(\lceil \frac{m}{2} \rceil - 1)} \epsilon^{-\gamma} \\ &\quad \cdot \sum_{l_1=0}^{k_1} \left[ \binom{k_1}{l_1} b_{\text{sm}}^{l_1} b_{\text{bg}}^{k_1-l_1} \right] \prod_{j=2}^m \sum_{l_j=0}^{k_j-1} \left[ \binom{k_j-1}{l_j} b_{\text{sm}}^{l_j} b_{\text{bg}}^{k_j-1-l_j} \right] \\ &= \epsilon^{-\gamma} \sum_{m=1}^{n+1} \sum_{\substack{k_1 + \dots + k_m = n \\ k_1 \in \mathbb{N}_0, k_2, \dots, k_m \in \mathbb{N}}} b_{\text{bg}}^{m-1} (\log(\epsilon^{1-2\gamma}))^{m-1} \epsilon^{(2\gamma-1)(\lceil \frac{m}{2} \rceil - 1)} (b_{\text{sm}} + b_{\text{bg}})^{k_1} \\ &\quad \cdot \prod_{j=2}^m (b_{\text{sm}} + b_{\text{bg}})^{k_j-1} \\ &= \epsilon^{-\gamma} \sum_{m=1}^{n+1} \sum_{\substack{k_1 + \dots + k_m = n \\ k_1 \in \mathbb{N}_0, k_2, \dots, k_m \in \mathbb{N}}} b_{\text{bg}}^{m-1} (\log(\epsilon^{1-2\gamma}))^{m-1} \epsilon^{(2\gamma-1)(\lceil \frac{m}{2} \rceil - 1)} (b_{\text{sm}} + b_{\text{bg}})^{n-(m-1)}. \end{aligned}$$

Given  $m$ , the number of configurations  $k_1 \in \mathbb{N}_0, k_2, \dots, k_m \in \mathbb{N}$  with  $k_1 + \dots + k_m = n$  is the number of arrangements of  $m - 1$  dividers and  $n - (m - 1)$  balls, which equals  $\binom{n}{m-1}$ . Since  $\lceil \frac{m}{2} \rceil - 1 \geq \frac{m-1}{2} - \frac{1}{2}$ , an application of the binomial theorem yields

$$\sum_{i_1, \dots, i_n \in \{1, \dots, 4\}} B_{i_n} \circ \dots \circ B_{i_1} g_1(\phi) \leq \epsilon^{-2\gamma+1/2} (b_{\text{bg}} \log(\epsilon^{1-2\gamma}) \epsilon^{\gamma-1/2} + b_{\text{bg}} + b_{\text{sm}})^n. \quad (2.8)$$

Combining (2.7) and (2.8), we conclude that for all  $\phi \in \Phi$

$$A^{2(n+1)} \mathbf{1}(\phi) \leq 4C_\epsilon \epsilon^{-2\gamma+1/2} b_{\text{bg}}^n (\log(\epsilon^{1-2\gamma}) \epsilon^{\gamma-1/2} + 1 + \frac{b_{\text{sm}}}{b_{\text{bg}}})^n.$$

Now (2.3) yields, for all  $\epsilon \in (0, 1)$ ,

$$\rho_\epsilon(A)^{-1} \geq \frac{2\gamma-1}{\sqrt{\beta(\gamma+\beta)}} \epsilon^{\gamma-1/2} \left( 1 + \log(\epsilon^{1-2\gamma}) \epsilon^{\gamma-1/2} + [\log(\epsilon^{1-2\gamma}) \epsilon^{\gamma-1/2}]^2 \right)^{-1/2}. \quad \square$$

The insight gained in the proof of the lower bound, enables us to ‘guess’ an approximating eigenfunction, which is the main ingredient in the proof of the upper bound.

*Proof of upper bound in Proposition 2.5 (b).* Let  $c_r := 1$  and  $c_\ell := \beta/(\gamma + \beta)$  and, for  $(\lambda, \alpha) \in \Phi$ , let

$$g_e(\lambda, \alpha) := c_\alpha \epsilon^\gamma e^{-\gamma\lambda} \mathbb{1}_{[\log \epsilon, \frac{\log \epsilon}{2}]}(\lambda) + \sqrt{\beta/(\gamma + \beta)} \epsilon^{1/2} e^{-(1-\gamma)\lambda} \mathbb{1}_{(\frac{\log \epsilon}{2}, 0]}(\lambda).$$

Notice that  $a_\alpha/c_\alpha = \gamma + \beta$  for  $\alpha \in \{\ell, r\}$ . Recall that we write  $|x|$  for the generation of a particle  $x$  in the IBP and  $\lambda(x)$  for its location. If  $(\lambda, \alpha) \in [\log \epsilon, \frac{\log \epsilon}{2}] \times \{\ell, r\}$ , then

$$\begin{aligned} Ag_e(\lambda, \alpha) &\geq E_{(\lambda, \alpha)} \left[ \sum_{\substack{|x|=1 \\ \lambda(x) > \frac{\log \epsilon}{2}}} g_e(\lambda(x), \ell) \right] = a_\alpha \sqrt{\beta/(\gamma + \beta)} \epsilon^{1/2} e^{-(1-\gamma)\lambda} \int_{-\lambda + \frac{\log \epsilon}{2}}^{-\lambda} e^{(2\gamma-1)t} dt \\ &= c_\alpha \frac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1} \epsilon^{1/2} e^{-\gamma\lambda} [1 - \epsilon^{\gamma-1/2}] = \frac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1} \epsilon^{-\gamma+1/2} [1 - \epsilon^{\gamma-1/2}] g_e(\lambda, \alpha). \end{aligned}$$

If  $(\lambda, \alpha) \in (\frac{\log \epsilon}{2}, 0] \times \{\ell, r\}$ , then

$$\begin{aligned} Ag_e(\lambda, \alpha) &\geq E_{(\lambda, \alpha)} \left[ \sum_{\substack{|x|=1 \\ \lambda(x) \leq \frac{\log \epsilon}{2}}} g_e(\lambda(x), r) \right] = \beta c_r \epsilon^\gamma e^{-\gamma\lambda} \int_{-\lambda + \log \epsilon}^{-\lambda + \frac{\log \epsilon}{2}} e^{(1-2\gamma)t} dt \\ &= \frac{\beta}{2\gamma-1} \epsilon^\gamma e^{-\gamma\lambda} e^{(1-2\gamma)(\log \epsilon - \lambda)} [1 - e^{-(1-2\gamma)\frac{\log \epsilon}{2}}] \\ &= \frac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1} \epsilon^{-\gamma+1/2} [1 - \epsilon^{\gamma-1/2}] g_e(\lambda, \alpha). \end{aligned}$$

By monotonicity of  $A$  this implies

$$\|A^n\| \geq \left( \frac{\sqrt{\beta(\gamma+\beta)}}{2\gamma-1} \epsilon^{1/2-\gamma} [1 - \epsilon^{\gamma-1/2}] \right)^n.$$

Taking the  $n$ -th root on both sides, an application of (2.3) yields the required bound for  $\rho_\epsilon(A)$ .  $\square$

### 2.3 A multitype branching process

In this section, we analyse the IBP and its relation to the associated operator  $A$ . We begin by providing properties of  $A$  in Section 2.3.1, and then use these properties to prove necessary and sufficient conditions for the multitype branching process to survive with positive probability in Section 2.3.2. Throughout, we use the notation introduced in Section 2.1 and write  $P_{\phi, p}$  for the distribution of the percolated IBP with retention probability  $p$  started with one particle of type  $\phi \in \Phi$ , abbreviating  $P_\phi := P_{\phi, 1}$ .

### 2.3.1 Proof of Proposition 2.2

**Lemma 2.6.** *For all nonnegative  $g \in C(\Phi)$  with  $g \not\equiv 0$ , we have  $\min_{\phi \in \Phi} A^2 g(\phi) > 0$ .*

*Proof.* If  $g \in C(\Phi)$ ,  $g \geq 0$ ,  $g \not\equiv 0$ , then there exist  $\log \epsilon \leq \lambda_1 < \lambda_2 \leq 0$  and  $\alpha_0 \in \{\ell, r\}$  such that  $g$  is strictly positive on  $[\lambda_1, \lambda_2] \times \{\alpha_0\}$ . Hence, it suffices to show that

$$\min_{\phi \in \Phi} P_\phi(\exists x: |x| = 2, \phi(x) \in [\lambda_1, \lambda_2] \times \{\alpha_0\}) > 0.$$

By the definition of the process any particle produces offspring in a given interval of positive length with, uniformly in the start type, strictly positive probability. The two steps allow the time needed to ensure that the relative position of the parent satisfies  $\alpha(x) = \alpha_0$ .  $\square$

**Lemma 2.7.** *The operator  $A: C(\Phi) \rightarrow C(\Phi)$  is compact.*

*Proof.* According to (2.4), we can write for  $g \in C(\Phi)$  and  $(\lambda, \alpha) \in \Phi$ ,

$$Ag(\lambda, \alpha) = \int_{\log \epsilon}^0 g(t, r) \kappa_\ell(\lambda, t) dt + \int_{\log \epsilon}^0 g(t, \ell) \kappa_r(\lambda, \alpha, t) dt,$$

with  $\kappa_\ell(\lambda, t) = \mathbb{1}_{[\log \epsilon, \lambda]}(t) \beta e^{(1-\gamma)(t-\lambda)}$  and  $\kappa_r(\lambda, \alpha, t) = \mathbb{1}_{[\lambda, 0]}(t) a_\alpha e^{\gamma(t-\lambda)}$ . Thus  $A$  can be written as the sum of two operators, which are both compact by the Arzelà-Ascoli theorem.  $\square$

We summarize some standard properties of compact, positive operators in the following proposition.

**Proposition 2.8.** *Let  $X$  be a complex Banach space and  $A: X \rightarrow X$  be a linear, compact and strictly positive operator.*

- (i) *The spectral radius of  $A$ ,  $\rho = \rho(A)$ , is a strictly positive eigenvalue of  $A$  with one dimensional eigenspace, generated by a strictly positive eigenvector  $\varphi$ . The eigenvalue  $\rho$  is also the spectral radius of adjoint  $A^*$  and the corresponding eigenspace is generated by a strictly positive eigenvector  $\nu_0$ . We rescale  $\varphi$  and  $\nu_0$  such that  $\|\varphi\| = 1$  and  $\nu_0(\varphi) = 1$  to make the choice unique.*
- (ii) *There exists  $\theta_0 \in [0, \rho)$  such that  $|\theta| \leq \theta_0$  for all  $\theta \in \sigma(A) \setminus \{\rho\}$ , where  $\sigma(A)$  is the spectrum of  $A$ .*
- (iii) *For any  $\theta > \theta_0$  and  $g \in X$ , we have  $A^n g = \rho^n \nu_0(g) \varphi + O(\theta^n)$ .*

*Proof.* Statements (i) and (ii) are immediate from the Krein-Rutman theorem, see Theorem 3.1.3 (ii) in [32], and the general form of the spectrum of compact operators. Statement (iii) then follows from the spectral decomposition of a compact operator on a complex Banach space. See for example [22] and there in particular Theorem 49.1 and Proposition 50.1.  $\square$

Now all results are collected to establish Proposition 2.2.

*Proof of Proposition 2.2.* Identity  $A_p = pA$  holds by definition and implies  $\rho_\epsilon(A_p) = p\rho_\epsilon(A)$ . Moreover, it is clear that it suffices to prove the first sentence of the statement for  $p = 1$ . Linearity is immediate from the definition, positivity was shown in Lemma 2.6 and compactness is the content of Lemma 2.7. The positive spectral radius follows immediately from Proposition 2.8 (i).  $\square$

### 2.3.2 Proof of Theorem 2.3

We start with a moment estimate for the total number of offspring of a particle. In the sequel, we write  $|\text{IBP}_n|$  for the number of particles in generation  $n$  of the IBP.

**Lemma 2.9.** *We have  $\sup_{\phi \in \Phi} E_\phi[|\text{IBP}_1|^2] < \infty$ .*

*Proof.* Let  $\Pi$ ,  $Z$  and  $\hat{Z}$  be independent realisations of the Poisson point process and the pure jump processes defined in Section 2.1. Let  $\phi = (\lambda, \alpha) \in \Phi$ . By the definition of the IBP,

$$|\text{IBP}_1| \stackrel{d}{=} \begin{cases} \Pi([\log \epsilon - \lambda, 0]) + Z_{-\lambda} & \text{if } \alpha = \ell, \\ \Pi([\log \epsilon - \lambda, 0]) + \hat{Z}_{-\lambda} & \text{if } \alpha = r, \end{cases}$$

where  $\stackrel{d}{=}$  denotes distributional equality. Since  $f$  is non-decreasing,  $\hat{Z}$  stochastically dominates  $Z$ . This implies that for all  $\phi \in \Phi$ ,

$$E_\phi[|\text{IBP}_1|^2] \leq 2 \left( \mathbb{E}[\Pi([\log \epsilon, 0])^2] + E[(\hat{Z}_{-\log \epsilon})^2] \right).$$

The first term on the right is finite because  $\Pi$  is a Poisson point process with finite intensity measure. The second summand was computed in Lemma 1.12 of [15] and found to be finite.  $\square$

The next result is a classical fact about branching processes. We give a proof since we could not find a reference for the result in sufficient generality. See Theorem III.11.2 in [21] for a special case.

**Lemma 2.10.** *For all  $p \in [0, 1]$ ,  $N \in \mathbb{N}$  and  $\phi \in \Phi$ ,*

$$P_{\phi,p}(1 \leq |\text{IBP}_n| \leq N \text{ infinitely often}) = 0.$$

*Proof.* We split the proof in two steps. First we show that  $\delta := \inf_{\phi \in \Phi} P_{\phi,p}(|\text{IBP}_1| = 0) > 0$ , then we conclude the statement from this result. By definition of the percolated IBP, for all  $(\lambda, \alpha) \in \Phi$ ,

$$\begin{aligned} P_{(\lambda,\alpha),p}(|\text{IBP}_1| = 0) &\geq P_{(\lambda,\alpha),1}(|\text{IBP}_1| = 0) \\ &= \begin{cases} P(\{\Pi([\log \epsilon - \lambda, 0]) = 0\} \cap \{Z_{-\lambda} = 0\}) & \text{if } \alpha = \ell \\ P(\{\Pi([\log \epsilon - \lambda, 0]) = 0\} \cap \{\hat{Z}_{-\lambda} = 0\}) & \text{if } \alpha = r \end{cases} \\ &\geq P(\Pi([\log \epsilon, 0]) = 0) P(\hat{Z}_{-\log \epsilon} = 0) > 0. \end{aligned}$$

Since the lower bound is independent of  $(\lambda, \alpha)$ , the claim  $\delta > 0$  is proved.

For the second step of the proof, we set  $p = 1$  to simplify notation. The proof for general  $p$  is identical. Fix  $N \in \mathbb{N}$ , set  $\tau_0 := 0$  and, for  $k \geq 1$ , let  $\tau_k := \inf\{n > \tau_{k-1} : |\text{IBP}_n| \in [1, N]\}$ , where  $\inf \emptyset := \infty$ . The strong Markov property implies, for all  $\phi \in \Phi$  and  $k \in \mathbb{N}$ ,

$$P_\phi(\tau_k < \infty) \leq P_\phi(\tau_1 < \infty) \sup_{\nu} P_\nu(\tau_1 < \infty)^{k-1},$$

where the supremum is over all counting measure  $\nu$  on  $\Phi$  such that  $\nu(\Phi) \in [1, N]$ . Under  $P_\nu$ ,  $\nu = \sum_{i=1}^n \delta_{\phi_i}$ , the branching process is started with  $n$  particles of types  $\phi_1, \dots, \phi_n$ . When all original ancestors have no offspring in the first generation, then the branching process suffers immediate extinction and  $\tau_1 = \infty$ . Hence, for all such  $\nu$ ,

$$P_\nu(\tau_1 < \infty) = 1 - P_\nu(\tau_1 = \infty) \leq 1 - P_\nu(|\text{IBP}_1| = 0) \leq 1 - \delta^{\nu(\Phi)} \leq 1 - \delta^N.$$

We conclude, for all  $\phi \in \Phi$

$$\begin{aligned} P_\phi(1 \leq |\text{IBP}_n| \leq N \text{ infinitely often}) &= \lim_{k \rightarrow \infty} P_\phi(\tau_k < \infty) \leq \lim_{k \rightarrow \infty} \sup_{\nu} P_\nu(\tau_1 < \infty)^{k-1} \\ &\leq \lim_{k \rightarrow \infty} (1 - \delta^N)^{k-1} = 0. \end{aligned} \quad \square$$



With Lemma 2.10 at hand, we can prove Theorem 2.3.

*Proof of Theorem 2.3.* Throughout the proof, we write  $\rho := \rho_\epsilon(A_p)$  and  $\varphi$  for the corresponding strictly positive eigenfunction with  $\|\varphi\| = 1$  from Proposition 2.8 (i). First suppose  $\rho \leq 1$ . By Lemma 2.10  $P_{\phi,p}(\lim_{n \rightarrow \infty} |\text{IBP}_n| \in \{0, \infty\}) = 1$ . By Proposition 2.8 (iii) the assumption  $\rho \leq 1$  implies that

$$E_{\phi,p}[|\text{IBP}_n|] = A_p^n \mathbf{1}(\phi) = \rho^n \nu_0(\mathbf{1}) \varphi(\phi) + o(1).$$

Hence,  $\sup_{n \in \mathbb{N}} E_{\phi,p}[|\text{IBP}_n|] < \infty$  and we conclude that  $\lim_{n \rightarrow \infty} |\text{IBP}_n| = 0$   $P_{\phi,p}$ -almost surely for all  $\phi \in \Phi$  and, therefore,  $\zeta^\epsilon(p) = 0$ .

Now suppose that  $\rho > 1$  and denote  $W_n = \frac{1}{\rho^n} \sum_{|x|=n} \varphi(\phi(x))$  for  $n \in \mathbb{N}$ . Then  $(W_n : n \in \mathbb{N})$  is under  $P_{\phi,p}$  a nonnegative martingale with respect to the filtration generated by the branching process. Hence,  $W := \lim_{n \rightarrow \infty} W_n$  exists almost surely. Given Lemma 2.9, Biggins and Kyprianou show in Theorem 1.1 of [5] that  $E_{\phi,p}[W] = 1$  and therefore,  $P_{\phi,p}(W > 0) > 0$ . This implies in particular that the branching process survives with positive probability irrespective of the start type.  $\square$

We now investigate continuity of the survival probability as a function of the attachment rule. For this purpose we emphasise dependence on  $f$  by adding it as an additional argument to several quantities. The result is used in the proof of Theorem 2.1 in Section 4 below.

**Lemma 2.11.** *Let  $p \in (0, 1]$ . Then  $\lim_{\delta \downarrow 0} \zeta^\epsilon(p, f - \delta) = \zeta^\epsilon(p, f)$ .*

*Proof.* Observe that there exists a natural coupling of the  $\text{IBP}(f)$  with the  $\text{IBP}(f - \delta)$  such that every particle in the  $\text{IBP}(f - \delta)$  is also present in the  $\text{IBP}(f)$  and, hence,  $\zeta^\epsilon(p, f - \delta)$  is increasing as  $\delta \downarrow 0$ . We can therefore assume that  $\zeta^\epsilon(p, f) > 0$ , that is  $\rho(f) := \rho_\epsilon(A_p, f) > 1$ , and by the continuity of  $A_p$  in the attachment rule, there exists  $\delta_0 > 0$  such that  $\rho_\epsilon(A_p, f - \delta_0) > 1$ . In the proof of Theorem 2.3 we have seen that this implies that the  $\text{IBP}(f - \delta_0)$  survives with positive probability, irrespective of the start type, and similar to Lemma 2.6, we conclude

$$\inf_{\phi \in \Phi} P_{\phi,p}(\text{IBP}(f - \delta_0) \text{ survives}) > 0. \quad (2.9)$$

Recall the definition of the martingale  $(W_n : n \in \mathbb{N})$  and its almost sure limit  $W$  from the proof of Theorem 2.3, which satisfies  $E_{\phi,p}[W] = 1$  and

$$W = \frac{1}{\rho(f)} \sum_{|x|=1} W(\phi(x)) \quad P_{\phi,p}\text{-almost surely,}$$

where, conditionally on the first generation,  $(W(\phi(x)) : |x| = 1)$  are independent copies of the random variable  $W$  under  $P_{\phi(x),p}$ . In particular,  $\phi \mapsto P_{\phi,p}(W = 0)$  is a fixed point of the operator  $Hg(\phi) = E_{\phi,p}[\prod_{|x|=1} g(\phi(x))]$  on the set of  $[0, 1]$ -valued, measurable functions. As the only  $[0, 1]$ -valued fixed points of  $H$  are the constant function 1 and the extinction probability  $\phi \mapsto P_{\phi,p}(\text{IBP}(f) \text{ dies out})$ , we deduce that  $W > 0$  almost surely on survival. Let  $c > 0$  and  $N \in \mathbb{N}$ . On the space of the coupling between  $\text{IBP}(f)$  and  $\text{IBP}(f - \delta)$ ,

$$\begin{aligned} \zeta^\epsilon(p, f) - \zeta^\epsilon(p, f - \delta) &= P(\text{IBP}(f) \text{ survives, IBP}(f - \delta) \text{ dies out}) \\ &\leq P(W \leq c, \text{IBP}(f) \text{ survives}) + P\left(W > c, \exists n \geq N : |\text{IBP}_n(f)| < \frac{c\rho(f)^n}{2 \max \varphi}\right) \\ &\quad + P\left(|\text{IBP}_n(f)| \geq \frac{c\rho(f)^n}{2 \max \varphi} \forall n \geq N, \text{IBP}(f - \delta) \text{ dies out}\right) \\ &=: \Theta_1(c) + \Theta_2(c, N) + \Theta_3(c, N, \delta). \end{aligned}$$

Since the offspring distribution of an individual particle is continuous in  $\delta$  uniformly on the type space, the probability that  $\text{IBP}(f)$  and  $\text{IBP}(f - \delta)$  agree until generation  $N$  tends to one as  $\delta \downarrow 0$ . On this event, when  $|\text{IBP}_N(f)| \geq C\rho(f)^N$  for some  $C > 0$ , then the probability that the  $\text{IBP}(f - \delta)$  subsequently dies out is bounded by

$$\sup_{\phi \in \Phi} P_{\phi,p}(\text{IBP}(f - \delta) \text{ dies out})^{\lceil C\rho(f)^N \rceil}.$$

By (2.9), this expression tends to zero as  $N \rightarrow \infty$  when  $\delta \leq \delta_0$ . Hence, for all  $c > 0$ ,

$$0 \leq \limsup_{\delta \downarrow 0} (\zeta^\epsilon(p, f) - \zeta^\epsilon(p, f - \delta)) \leq \Theta_1(c) + \limsup_{N \rightarrow \infty} \Theta_2(c, N).$$

On the event  $\{W > c\} \cap \{W_n \rightarrow W\}$ , there is a finite stopping time  $N_0$  such that  $W_n \geq W/2$  for all  $n \geq N_0$  and we deduce that  $\rho(f)^{-n} |\text{IBP}_n(f)| \geq W_n / \max \varphi \geq c/(2 \max \varphi)$ . Since  $W_n$  converges to  $W$  almost surely, we conclude that  $\lim_{N \rightarrow \infty} \Theta_2(c, N) = 0$ . Finally,  $\Theta_1(c)$  tends to zero as  $c \downarrow 0$  because  $W$  is positive on the event of survival.  $\square$

### 3 The topology of the damaged graph

We investigate the empirical indegree distribution and maximal indegree of the damaged network in Section 3.1, and typical distances in Section 3.2.

#### 3.1 Degrees

The following lemma formalises basic facts about the indegrees  $\mathcal{Z}[m, n]$ .

**Lemma 3.1.** *For given  $n \in \mathbb{N}$ , the random variables  $(\mathcal{Z}[m, n] : m \leq n)$  are independent. Fix  $\hat{m} \in \mathbb{N}$  and let  $(\mathcal{Z}^m[\hat{m}, n] : m \leq n)$ , be independent copies of the random variable  $\mathcal{Z}[\hat{m}, n]$ .*

- (i) *There exists a coupling between  $(\mathcal{Z}[m, n] : 1 \leq m \leq \hat{m})$  and  $(\mathcal{Z}^m[\hat{m}, n] : 1 \leq m \leq \hat{m})$  such that*

$$\mathcal{Z}[m, n] \geq \mathcal{Z}^m[\hat{m}, n] \quad \text{for all } 1 \leq m \leq \hat{m}.$$

- (ii) *There exists a coupling between  $(\mathcal{Z}[m, n] : \hat{m} \leq m \leq n)$  and  $(\mathcal{Z}^m[\hat{m}, n] : \hat{m} \leq m \leq n)$  such that*

$$\mathcal{Z}[m, n] \leq \mathcal{Z}^m[\hat{m}, n] \quad \text{for all } \hat{m} \leq m \leq n.$$

*Proof.* The independence of  $(\mathcal{Z}[m, n] : m \leq n)$  is immediate from the network construction. Consequently, to prove (i) and (ii) it suffices to couple  $\mathcal{Z}[m, n]$  and  $\mathcal{Z}^m[\hat{m}, n]$  for fixed  $1 \leq m \leq \hat{m} \leq n$  in such a way that  $\mathcal{Z}[m, n] \geq \mathcal{Z}^m[\hat{m}, n]$ . Equivalently, we show that  $\mathcal{Z}[m, n]$  stochastically dominates  $\mathcal{Z}[\hat{m}, n]$ . Let  $\mathcal{Y}^m = (\mathcal{Y}_l^m : l \in \mathbb{N}_0)$  be the Markov process given by  $\mathcal{Y}_l^m = \mathcal{Z}[m, m + l]$ . Then  $\mathcal{Y}_0^m = 0$  and, for all  $l, k \in \mathbb{N}_0$  with  $k \leq l$ ,

$$\mathbb{P}(\mathcal{Y}_{l+1}^m = k + 1 | \mathcal{Y}_l^m = k) = 1 - \mathbb{P}(\mathcal{Y}_{l+1}^m = k | \mathcal{Y}_l^m = k) = \frac{f(k)}{m + l}$$

is decreasing in  $m$ . Hence, in every step, the probability that  $\mathcal{Y}^m$  jumps is at least the probability that  $\mathcal{Y}^{\hat{m}}$  jumps. Since  $\mathcal{Z}[m, n] = \mathcal{Y}_{n-m}^m$ ,  $\mathcal{Z}[\hat{m}, n] = \mathcal{Y}_{n-\hat{m}}^{\hat{m}}$  and  $n - m \geq n - \hat{m}$ , the claim is established.  $\square$

Our goal is to determine the asymptotic behaviour of  $\max_{m \in \mathbb{V}_n^\epsilon} \mathcal{Z}[m, n]$  and

$$X_{\geq k}^\epsilon(n) = \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{m=\lfloor \epsilon n \rfloor + 1}^n \mathbb{1}_{\{k, k+1, \dots\}}(\mathcal{Z}[m, n]).$$

Lemma 3.1 allows us to replace the independent random variables in these sequences by groups of independent and identically distributed random variables.

Dereich and Mörters observe in [14], see for example Corollary 4.3, that the indegrees in network  $(G_n : n \in \mathbb{N})$  are closely related to the pure jump process  $(Z_t)_{t \geq 0}$ . Since the indegrees are not altered by the targeted attack, the same holds in the damaged network  $(G_n^\epsilon : n \in \mathbb{N})$ . We now explain this connection. Let  $\psi(k) := \sum_{j=1}^{k-1} j^{-1}$  for all  $k \in \mathbb{N}$ , which we consider as a time change, mapping real time epochs  $k$  to an ‘artificial time’  $\psi(k)$ . The artificial time spent by the process  $\mathcal{Z}[m, \cdot]$  in state  $i$  is

$$T_m[i] := \sup \left\{ \sum_{j=k}^l \frac{1}{j} : \mathcal{Z}[m, k] = i = \mathcal{Z}[m, l] \right\}.$$

Let  $k \in \mathbb{N}_0$  and  $n_k$  the last real time that  $\mathcal{Z}[m, \cdot]$  spends in  $k$ , that is,  $\mathcal{Z}[m, n_k] = k$ ,  $\mathcal{Z}[m, n_k + 1] = k + 1$ . Then  $\sum_{i=0}^k T_m[i] = \sum_{j=m}^{n_k} j^{-1} = \psi(n_k + 1) - \psi(m)$ . In particular,

$$\mathcal{Z}[m, n] \leq k \Leftrightarrow \sum_{i=0}^k T_m[i] \geq \psi(n + 1) - \psi(m). \quad (3.1)$$

By definition, there exists a sequence of independent random variables  $(T[i] : i \in \mathbb{N}_0)$  such that  $T[i]$  is exponentially distributed with mean  $1/f(i)$  and

$$Z_t \leq k \Leftrightarrow \sum_{i=0}^k T[i] > t. \quad (3.2)$$

The next lemma provides a coupling between the artificial times  $T_m[i]$  and the exponential times  $T[i]$ . In combination with (3.1) and (3.2) this allows us to study the jump process  $(Z_t : t \geq 0)$  instead of the involved dynamics of the indegree process  $\mathcal{Z}[m, \cdot]$ . The proof of the lemma is identical to the proof of Lemma 4.1 in [14] and is omitted. However, the argument behind the result was sketched in the paragraph of (1.9). We denote by  $\tau_{m,i} = \inf\{\psi(k) : \mathcal{Z}[m, k] = i\}$  the artificial first entrance time of  $\mathcal{Z}[m, \cdot]$  into state  $i$ . If  $\tau_{m,i} = \psi(k)$ , we write  $\Delta\tau_{m,i} = k^{-1}$ .

**Lemma 3.2.** *There exists a constant  $\eta > 0$  such that for all  $m \in \mathbb{N}$  there is a coupling such that, for all  $i \in \mathbb{N}_0$  with  $f(i)\Delta\tau_{m,i} \leq \frac{1}{2}$ ,*

$$T[i] - \eta f(i) \Delta\tau_{m,i} \leq T_m[i] \leq T[i] + \Delta\tau_{m,i} \quad \text{almost surely,}$$

*and the random variables  $((T[i], T_m[i]) : i \in \mathbb{N}_0)$  are independent.*

By definition  $\Delta\tau_{m,i} \leq m^{-1}$ . Hence, Lemma 3.2 yields a coupling such that when  $f(k)/m \leq \frac{1}{2}$ , then

$$\sum_{i=0}^k T[i] - k\eta f(k)/m \leq \sum_{i=0}^k T_m[i] \leq \sum_{i=0}^k T[i] + k/m.$$

In particular, for  $f(k)/m \leq \frac{1}{2}$ , the equivalence (3.1) implies

$$\begin{aligned} P\left(\sum_{i=0}^k T[i] \geq \psi(n + 1) - \psi(m) + \eta k f(k)/m\right) &\leq \mathbb{P}(\mathcal{Z}[m, n] \leq k) \\ &\leq P\left(\sum_{i=0}^k T[i] \geq \psi(n + 1) - \psi(m) + k/m\right). \end{aligned}$$

If  $m \leq n$  and  $m - \vartheta n = O(1)$  for some  $\vartheta \in (0, 1]$ , then  $\psi(n+1) - \psi(m) = \sum_{j=m}^n j^{-1} = -\log \vartheta + o(1)$ . Hence, for  $m \leq n$ ,  $m - \vartheta n = O(1)$  and  $k = O(\log n)$ ,

$$\mathbb{P}(\mathcal{Z}[m, n] \leq k) = P\left(\sum_{i=0}^k T[i] \geq -\log \vartheta + o(1)\right), \quad (3.3)$$

where the random null sequence  $o(1)$  is bounded by a deterministic null sequence of order  $O((\log n)^2/n)$ .

We proceed by estimating the distribution function of  $\sum_{i=0}^k T[i]$ . The following identity for the incomplete beta function will be of use.

**Lemma 3.3.** *Let  $a > 0$ ,  $c > 0$  and  $k \in \mathbb{N}_0$ . Then*

$$\sum_{i=0}^k \binom{k}{i} \frac{(-a)^i}{i+c} = a^{-c} \int_0^a x^{c-1} (1-x)^k dx.$$

*Proof.* Denote the left-hand side by  $\theta(k, a, c)$ . For  $x > 0$ , we have

$$\frac{\partial}{\partial x} [x^c \theta(k, x, c)] = \frac{\partial}{\partial x} \left[ \sum_{i=0}^k \binom{k}{i} (-1)^i \frac{x^{i+c}}{i+c} \right] = \sum_{i=0}^k \binom{k}{i} (-1)^i x^{i+c-1} = x^{c-1} (1-x)^k.$$

Integrating both sides between 0 and  $a$  and dividing by  $a^c$ , we obtain the claim.  $\square$

**Lemma 3.4.** *For  $k \in \mathbb{N}_0$  and  $t \geq 0$ ,*

$$P(Z_t \geq k+1) = P\left(\sum_{i=0}^k T[i] \leq t\right) = B\left(k+1, \frac{\beta}{\gamma}\right)^{-1} \int_{e^{-\gamma t}}^1 x^{\frac{\beta}{\gamma}-1} (1-x)^k dx. \quad (3.4)$$

*Proof.* Let  $k \in \mathbb{N}_0$ . The probability density for  $\sum_{i=0}^k T[i]$  is given by (see for example Problem 12, Chapter 1 in [19])

$$t \mapsto \sum_{i=0}^k \frac{\prod_{j=0, j \neq i}^k f(j)}{\prod_{j=0, j \neq i}^k (f(j) - f(i))} f(i) e^{-f(i)t} \mathbb{1}_{[0, \infty)}(t).$$

Using  $f(j) = \gamma j + \beta$ , we can rewrite for all  $i \in \{0, \dots, k\}$ ,

$$\frac{\prod_{j=0, j \neq i}^k f(j)}{\prod_{j=0, j \neq i}^k (f(j) - f(i))} = \frac{\frac{\beta}{f(i)} \frac{k!}{i!} \prod_{j=1}^k f(j)}{\gamma^k (-1)^i i! (k-i)!} = \frac{\beta}{\gamma} \left( \prod_{j=1}^k \frac{f(j)}{\gamma j} \right) \binom{k}{i} \frac{(-1)^i}{i + \frac{\beta}{\gamma}}.$$

We obtain

$$P\left(\sum_{i=0}^k T[i] \leq t\right) = \frac{\beta}{\gamma} \left( \prod_{j=1}^k \frac{f(j)}{\gamma j} \right) \sum_{i=0}^k \binom{k}{i} \frac{(-1)^i}{i + \frac{\beta}{\gamma}} (1 - e^{-\beta t} e^{-\gamma i t}).$$

The factor in front of the sum equals  $B(k+1, \frac{\beta}{\gamma})^{-1}$ . Thus, it remains to show that the sum agrees with the integral in (3.4). Using Lemma 3.3, we derive

$$\begin{aligned} & \sum_{i=0}^k \binom{k}{i} \frac{(-1)^i}{i + \frac{\beta}{\gamma}} (1 - e^{-\beta t} e^{-\gamma i t}) \\ &= \int_0^1 x^{\frac{\beta}{\gamma}-1} (1-x)^k dx - e^{-\beta t} (e^{-\gamma t})^{-\frac{\beta}{\gamma}} \int_0^{e^{-\gamma t}} x^{\frac{\beta}{\gamma}-1} (1-x)^k dx \\ &= \int_{e^{-\gamma t}}^1 x^{\frac{\beta}{\gamma}-1} (1-x)^k dx. \end{aligned} \quad \square$$

Now everything is prepared for the proof of Theorem 1.2.

**Proposition 3.5.** *For all  $k \in \mathbb{N}_0$*

$$\mathbb{E}[X_{\geq k+1}^\epsilon(n)] \rightarrow \int_\epsilon^1 \frac{1}{1-\epsilon} B(k+1, \frac{\beta}{\gamma})^{-1} \int_{y^\gamma}^1 x^{\frac{\beta}{\gamma}-1} (1-x)^k dx dy \quad \text{as } n \rightarrow \infty.$$

*Proof.* Let  $\delta > 0$ ,  $\Delta = \lfloor \delta \epsilon n \rfloor$ ,  $N = 1 + \lfloor \frac{n - \lfloor \epsilon n \rfloor}{\Delta} \rfloor$ ,  $m_j = \lfloor \epsilon n \rfloor + 1 + j\Delta$  for  $j = 0, \dots, N-1$ ,  $m_N = n+1$ ,  $\Delta_j = \Delta$  for  $j = 1, \dots, N-1$  and  $\Delta_N = m_N - m_{N-1} \in [0, \Delta]$ . Combining Lemma 3.1 with (3.3), we obtain

$$\begin{aligned} \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{m=\lfloor \epsilon n \rfloor+1}^n \mathbb{P}(\mathcal{Z}[m, n] > k) &\leq \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{j=0}^{N-1} \Delta_{j+1} \mathbb{P}(\mathcal{Z}[m_j, n] > k) \\ &\leq \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{j=0}^{N-1} \delta \epsilon n P\left(\sum_{i=0}^k T[i] \leq -\log(\epsilon + j\delta\epsilon) + o(1)\right), \end{aligned}$$

and

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{m=\lfloor \epsilon n \rfloor+1}^n \mathbb{P}(\mathcal{Z}[m, n] > k) \\ \leq \sum_{j=0}^{\lceil \frac{1-\epsilon}{\delta\epsilon} \rceil} \frac{\delta\epsilon}{1-\epsilon} P\left(\sum_{i=0}^k T[i] \leq -\log\left(\epsilon + j\frac{\delta\epsilon}{1-\epsilon}(1-\epsilon)\right)\right). \end{aligned}$$

Taking  $\delta \rightarrow 0$  on the right-hand side, we conclude

$$\limsup_{n \rightarrow \infty} \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{m=\lfloor \epsilon n \rfloor+1}^n \mathbb{P}(\mathcal{Z}[m, n] > k) \leq \int_0^1 P\left(\sum_{i=0}^k T[i] \leq -\log(\epsilon + y(1-\epsilon))\right) dy. \quad (3.5)$$

Similarly,

$$\begin{aligned} \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{m=\lfloor \epsilon n \rfloor+1}^n \mathbb{P}(\mathcal{Z}[m, n] > k) &\geq \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{j=1}^N \Delta_j \mathbb{P}(\mathcal{Z}[m_j - 1, n] > k) \\ &\geq \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{j=1}^{N-1} \Delta_j P\left(\sum_{i=0}^k T[i] \leq -\log(\epsilon + j\delta\epsilon) + o(1)\right), \end{aligned}$$

and as above we see that  $\liminf$  satisfies the reverse inequality in (3.5). Lemma 3.4 yields the claim.  $\square$

Dereich and Mörters (pp 1238–1239 in [14]) give a simple argument based on Chernoff's inequality to upgrade the convergence of the expected empirical degree distribution to convergence of the empirical degree distribution itself. The proof remains valid for the damaged network and is therefore omitted.

To establish the claimed tail behaviour of  $\mu$ , we consider the large  $k$  asymptotics of  $P(Z_t \geq k+1)$  in (3.4). By Stirling's formula,  $B(k+1, \beta/\gamma)^{-1} \asymp k^{\beta/\gamma}$ , where we write  $a(k) \asymp b(k)$  if there exist constants  $0 < c \leq C < \infty$  such that  $ca(k) \leq b(k) \leq Ca(k)$  for all large  $k$ . Moreover,

$$\begin{aligned} \frac{1}{1-\epsilon} \int_\epsilon^1 \int_{y^\gamma}^1 x^{\frac{\beta}{\gamma}-1} (1-x)^k dx dy &\asymp \int_\epsilon^1 \int_{y^\gamma}^1 (1-x)^k dx dy \\ &= \frac{1}{k+1} \int_\epsilon^1 (1-y^\gamma)^{k+1} dy \asymp \frac{(1-\epsilon^\gamma)^{k+1}}{(k+1)^2}. \end{aligned}$$

In the first estimate we used that  $x^{\frac{\beta}{\gamma}-1}$  is bounded from zero and infinity; in the second we employed Laplace's method (see for example Section 3.5 of [28]). In particular,  $\mu^\epsilon$  has the stated tails. To complete the proof of Theorem 1.2, it remains to derive the asymptotic behaviour of the maximal indegree. The statement follows from the next two lemmas.

**Lemma 3.6** (Upper bound). *Let  $c > -\frac{1}{\log(1-\epsilon^\gamma)}$ . Then,*

$$\mathbb{P}\left(\max_{m \in V_n^\epsilon} \mathcal{Z}[m, n] \leq c \log n\right) \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

*Proof.* Write  $k_n = \lfloor c \log n \rfloor$ ,  $\underline{m} = \lfloor \epsilon n \rfloor + 1$  and  $\Delta = n - \lfloor \epsilon n \rfloor$ . Moreover, let  $\mathcal{Z}^m[\underline{m}, n]$ ,  $m \leq n$ , be independent copies of  $\mathcal{Z}[\underline{m}, n]$ . Lemma 3.1 and (3.3) yield

$$\begin{aligned} \mathbb{P}\left(\max_{m \in V_n^\epsilon} \mathcal{Z}[m, n] \leq c \log n\right) &\geq \mathbb{P}\left(\max_{m \in V_n^\epsilon} \mathcal{Z}^m[\underline{m}, n] \leq k_n\right) = \mathbb{P}\left(\mathcal{Z}[\underline{m}, n] \leq k_n\right)^\Delta \\ &= P\left(\sum_{i=0}^{k_n} T[i] \geq -\log \epsilon + o(1)\right)^\Delta \\ &= \exp\left(-\Delta P\left(\sum_{i=0}^{k_n} T[i] \leq -\log \epsilon + o(1)\right)(1 + o(1))\right), \end{aligned} \quad (3.6)$$

using a Taylor expansion in the last equality. As above, uniformly for  $t$  in compact subintervals of  $(0, \infty)$ ,

$$\int_{e^{-\gamma t}}^1 x^{\frac{\beta}{\gamma}-1} (1-x)^k dx \asymp \int_{e^{-\gamma t}}^1 (1-x)^k dx \asymp \exp\left(k \log(1 - e^{-\gamma t}) - \log k\right).$$

Thus, Lemma 3.4 and Stirling's formula yield for  $\vartheta \in (0, 1)$  and  $t = -\log \vartheta + o(1)$ , as  $k \rightarrow \infty$ ,

$$P\left(\sum_{i=0}^k T[i] \leq t\right) \asymp \exp\left(k \log(1 - e^{-\gamma t}) + \left(\frac{\beta}{\gamma} - 1\right) \log k\right) = \exp\left(k \log(1 - \vartheta^\gamma)(1 + o(1))\right). \quad (3.7)$$

Using this estimate for  $k = k_n$  and  $\vartheta = \epsilon$ , the exponent on the right-hand side of (3.6) tends to zero as  $n \rightarrow \infty$  if  $c > -1/\log(1 - \epsilon^\gamma)$ .  $\square$

**Lemma 3.7** (Lower bound). *Let  $c < -\frac{1}{\log(1-\epsilon^\gamma)}$ . Then,*

$$\mathbb{P}\left(\max_{m \in V_n^\epsilon} \mathcal{Z}[m, n] \leq c \log n\right) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

*Proof.* The idea of the proof is to restrict the maximum to an arbitrarily small proportion of the oldest vertices. Let  $\delta > 0$  and write  $k_n := \lfloor c \log n \rfloor$ ,  $\Delta = \lfloor \delta \epsilon n \rfloor$  and  $\bar{m} = \lfloor \epsilon n \rfloor + \Delta$ . Moreover, let  $\mathcal{Z}^m[\bar{m}, n]$ ,  $m \leq n$ , be independent copies of  $\mathcal{Z}[\bar{m}, n]$ . According to Lemma 3.1 there is a coupling such that

$$\max_{m \in V_n^\epsilon} \mathcal{Z}[m, n] \geq \max_{m=\lfloor \epsilon n \rfloor+1, \dots, \lfloor \epsilon n \rfloor+\Delta} \mathcal{Z}[m, n] \geq \max_{m=\lfloor \epsilon n \rfloor+1, \dots, \lfloor \epsilon n \rfloor+\Delta} \mathcal{Z}^m[\bar{m}, n].$$

Arguing as in (3.6), (3.3) yields

$$\mathbb{P}\left(\max_{m \in V_n^\epsilon} \mathcal{Z}[m, n] \leq c \log n\right) \leq \exp\left(-\Delta P\left(\sum_{i=0}^{k_n} T[i] \leq -\log(\epsilon(1+\delta)) + o(1)\right)(1 + o(1))\right).$$

Now (3.7) with  $\vartheta = \epsilon(1+\delta)$  implies that the exponent on the right-hand side tends to  $-\infty$  if  $c < -1/\log(1 - (\epsilon(1+\delta))^\gamma)$ . Since  $\delta$  was arbitrary, the claim is established.  $\square$

### 3.2 Distances

In this section, we study the typical distance between two uniformly chosen vertices in  $C_n^\epsilon$  and prove Theorem 1.3. We write  $\Delta Z[m, n] = Z[m, n+1] - Z[m, n]$  and, for  $m \geq n$ ,  $Z[m, n] = 0$ . In the graph, the indegree of vertex  $m$  at time  $m$  is zero by definition, but we will also use the distribution of the process  $(Z[m, n]: n \geq m)$  for different initial values. Formally, the evolution of  $Z[m, \cdot]$  with initial value  $k$  is obtained by using the attachment rule  $g(l) := f(k + l)$  and we denote its distribution by  $\mathbb{P}^k$ , using  $\mathbb{E}^k$  for the corresponding expectation; we abbreviate  $\mathbb{P} := \mathbb{P}^0$ ,  $\mathbb{E} := \mathbb{E}^0$ . We further write  $\hat{n} := \inf\{n \in \mathbb{N}: f(n)/n \leq 1\} \vee 2$ . Note that  $\gamma < 1$  implies  $\hat{n} \in \mathbb{N}$ . We observe some facts about the indegree distribution. These are adaptations of results in [15].

**Lemma 3.8** (Lemma 2.7 in [15]). *For all  $k \in \mathbb{N}_0$  and  $m, n \in \mathbb{N}$  with  $k \leq m$ ,  $\hat{n} \leq m \leq n$ ,*

$$\mathbb{P}^k(\Delta Z[m, n] = 1) \leq \frac{f(k)}{(m-1)^{\gamma} n^{1-\gamma}}. \quad (3.8)$$

*Proof.* Observe that  $(f(Z[m, n]) \prod_{j=m}^{n-1} \frac{1}{1+\gamma/j} : n \geq m)$  is a martingale and therefore

$$\mathbb{P}^k(\Delta Z[m, n] = 1) = \mathbb{E}^k \left[ \frac{f(Z[m, n])}{n} \right] = \frac{f(k)}{n} \prod_{j=m}^{n-1} (1 + \gamma/j) \leq \frac{f(k)}{(m-1)^{\gamma} n^{1-\gamma}}. \quad \square$$

**Lemma 3.9** (Lemma 2.10 in [15]). *For all  $k \in \mathbb{N}_0$ ,  $m, m' \in \mathbb{N}$ ,  $\hat{n} \leq m \leq m'$ ,  $k \leq m$ , there exists a coupling of the process  $(Z[m, n]: n \geq m)$  under the conditional probability  $\mathbb{P}^k(\cdot | \Delta Z[m, m'] = 1)$  and the process  $(Z[m, n]: n \geq m)$  under  $\mathbb{P}^{k+1}$  such that, apart from time  $m'$ , the jump times of the first process are a subset of the jump times of the latter.*

The proof of the lemma is similar to the proof of Lemma 2.10 in [15] and we omit it. After these preliminary results, we now begin our analysis of typical distances in the network  $(G_n^\epsilon: n \in \mathbb{N})$ . Recall, that for this type of questions, we consider  $G_n^\epsilon$  to be an undirected graph. For  $v, w \in V_n^\epsilon$  and  $h \in \mathbb{N}_0$ , let

$$\mathcal{S}_h(v, w) := \{(v_0, \dots, v_h): v_i \in V_n^\epsilon, v_i \neq v_j \text{ for } i \neq j, v_0 = v, v_h = w\}$$

be the set of all self-avoiding paths of length  $h$  between  $v$  and  $w$ , and let  $\mathcal{S}_h(v) = \{p: p \in \mathcal{S}_h(v, w) \text{ for some } w \in V_n^\epsilon\}$  the set of all self-avoiding paths of length  $h$  starting in  $v$ .

**Definition 3.10.** Let  $\theta \in (0, \infty)$  and  $G = (V, E)$  be an undirected graph with  $V \subseteq \mathbb{N}$ . A self-avoiding path  $p = (v_0, \dots, v_h)$  in  $G$  is  $\theta$ -admissible (or admissible) if, for all  $i \in \{1, \dots, h\}$ , we have  $\{v_{i-1}, v_i\} \in E$  and

$$|\{w \in V: v_{i-1} < w \leq v_i, \{v_{i-1}, w\} \in E\}| \leq \theta. \quad (3.9)$$

Note that (3.9) is automatically satisfied if  $v_i < v_{i-1}$ . In the graph  $G_n^\epsilon$ , condition (3.9) can be written as  $Z[v_{i-1}, v_i] \leq \theta$ . We further denote, for  $v, w \in V_n^\epsilon$ ,  $h \in \mathbb{N}_0$  and  $\theta \in (0, \infty)$ ,

$$N_h^\theta(v, w) := |\{p \in \mathcal{S}_h(v, w): p \text{ is } \theta\text{-admissible in } G_n^\epsilon\}|,$$

$$N_h^\theta(v) := |\{p \in \mathcal{S}_h(v): p \text{ is } \theta\text{-admissible in } G_n^\epsilon\}|,$$

and, for  $h > 0$ ,  $N_{\leq h}^\theta(v) = \sum_{k=0}^{[h]} N_k^\theta(v)$ ,  $N_{\leq h}^\theta(v, w) = \sum_{k=0}^{[h]} N_k^\theta(v, w)$ . The dependence of  $\mathcal{S}_h(v, w)$ ,  $\mathcal{S}_h(v)$ ,  $N_h^\theta(v, w)$  etc. on  $n$  is suppressed in the notation, but it will always be clear from the context which graph is considered. We write  $\text{IBP}^\epsilon(f)$  for the idealized branching process with type space  $[\log \epsilon, 0] \times \{\ell, r\}$  generated with attachment rule  $f$  if we want to emphasize  $f$  and  $\epsilon$ . The proof of the following lemma is deferred to Section 4.3.

**Lemma 3.11.** Let  $\delta > 0$  such that  $\gamma(1 + \delta) < 1$ ,  $\epsilon \in (0, \epsilon)$ , and  $(\theta_n : n \in \mathbb{N})$  a sequence of positive numbers with  $\theta_n = o(n)$ . For all sufficiently large  $n$ ,  $v_0 \in V_n^\epsilon$ , and  $h \in \mathbb{N}_0$ ,

$$\mathbb{E}[N_h^{\theta_n}(v_0)] \leq E_{(s_n(v_0), \ell)}[|\text{IBP}_h^\epsilon((1 + \delta)f)|] \quad \text{where } s_n(v_0) := -\sum_{j=v_0}^{n-1} \frac{1}{j}.$$

We are now in the position to prove Theorem 1.3. The distance between two vertices  $v, w \in V_n^\epsilon$  in different components of  $G_n^\epsilon$  is defined to be infinite.

*Proof of Theorem 1.3.* Let  $v, w \in V_n^\epsilon$ ,  $h \in \mathbb{N}$ . With  $\theta_n := (\log n)^2$ , (1.5) yields

$$\begin{aligned} \mathbb{P}(d_{G_n^\epsilon}(v, w) \leq h) &\leq \mathbb{P}\left(d_{G_n^\epsilon}(v, w) \leq h, \max_{m \in V_n^\epsilon} \mathcal{Z}[m, n] \leq \theta_n\right) + \mathbb{P}\left(\max_{m \in V_n^\epsilon} \mathcal{Z}[m, n] > \theta_n\right) \\ &\leq \mathbb{P}(N_{\leq h}^{\theta_n}(v, w) \geq 1) + o(1), \end{aligned} \quad (3.10)$$

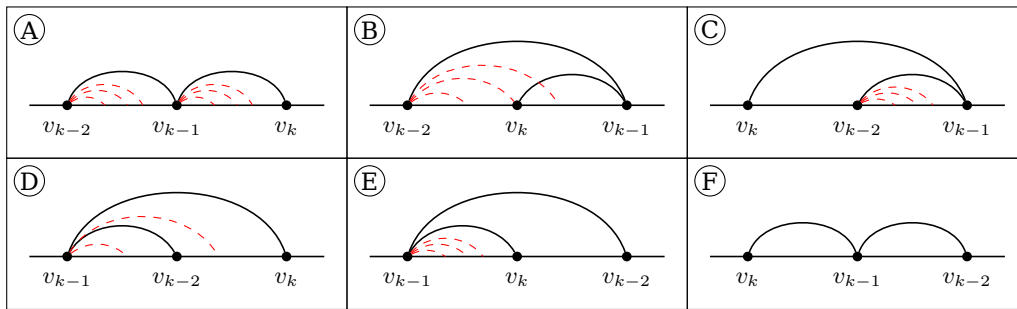
where the error bound is uniform in  $v, w$  and  $h$ . Markov's inequality yields, for every  $v, w \in V_n^\epsilon$  with  $v \neq w$  and for every  $h \in \mathbb{N}$ ,

$$\mathbb{P}(N_{\leq h}^{\theta_n}(v, w) \geq 1) \leq \mathbb{E}[N_{\leq h}^{\theta_n}(v, w)] = \sum_{k=1}^h \sum_{p \in \mathcal{S}_k(v, w)} \mathbb{P}(p \text{ is } \theta_n\text{-admissible in } G_n^\epsilon). \quad (3.11)$$

We write  $w_i^+ := v_{i-1} \vee v_i$ ,  $w_i^- := v_{i-1} \wedge v_i$  and  $\mathcal{E}_i := \{\Delta \mathcal{Z}[w_i^-, w_i^+ - 1] = 1, \mathcal{Z}[v_{i-1}, v_i] \leq \theta_n\}$  for every  $i \in \{1, \dots, k\}$ . Then we have

$$\begin{aligned} \mathbb{P}(p \text{ is } \theta_n\text{-admissible in } G_n^\epsilon) &= \mathbb{P}\left(\bigcap_{i=1}^k \{\Delta \mathcal{Z}[w_i^-, w_i^+ - 1] = 1, \mathcal{Z}[v_{i-1}, v_i] \leq \theta_n\}\right) \\ &= \mathbb{P}\left(\mathcal{E}_k \mid \bigcap_{i=1}^{k-1} \mathcal{E}_i\right) \mathbb{P}((v_0, \dots, v_{k-1}) \text{ is } \theta_n\text{-admissible in } G_n^\epsilon). \end{aligned} \quad (3.12)$$

To estimate the probability  $\mathbb{P}(\mathcal{E}_k \mid \bigcap_{i=1}^{k-1} \mathcal{E}_i)$ , we first note that the only edge in the self-avoiding path  $p$  on whose presence the event  $\{v_{k-1}, v_k\} \in E_n^\epsilon$  can depend is  $\{v_{k-2}, v_{k-1}\}$ . The possible arrangements of these two edges are sketched in Figure 4.



**Figure 4:** Possible interactions of two edges on a self-avoiding path. The red, dashed edges have to be considered to decide if the number of right-neighbours is small enough to declare the path admissible.

When  $v_{k-2} < v_{k-1}$  (cases A,B,C in Figure 4), then we, in addition, have knowledge of edges whose left vertex is  $v_{k-2}$ . However, these are always independent of  $\{v_{k-1}, v_k\}$ . If  $v_{k-1} < v_k$  (cases A,D,E in Figure 4), then event  $\mathcal{E}_k$  requires that  $\mathcal{Z}[v_{k-1}, v_k] \leq \theta_n$ . Since edges with left vertex  $v_{k-1}$  depend only on edges whose left vertex is also  $v_{k-1}$ , the only



relevant conditioning occurs in cases D and E in Figure 4 by requiring  $\{v_{k-1}, v_{k-2}\}$  to be present. We deduce

$$\begin{aligned} & \mathbb{P}\left(\mathcal{E}_k \middle| \bigcap_{i=1}^{k-1} \mathcal{E}_i\right) \\ &= \begin{cases} \mathbb{P}(\Delta \mathcal{Z}[w_k^-, w_k^+ - 1] = 1, \mathcal{Z}[v_{k-1}, v_k] \leq \theta_n) & \text{in A,B,C and F,} \\ \mathbb{P}(\Delta \mathcal{Z}[v_{k-1}, v_k - 1] = 1, \mathcal{Z}[v_{k-1}, v_k] \leq \theta_n | \Delta \mathcal{Z}[v_{k-1}, v_{k-2} - 1] = 1) & \text{in D and E.} \end{cases} \end{aligned}$$

Using Lemma 3.9 and (3.8), we can bound the probability in both cases by  $f(1)/(\epsilon n)$ . Combining this estimate with (3.11) and (3.12), we obtain

$$\mathbb{P}(N_{\leq h}^{\theta_n}(v, w) \geq 1) \leq \sum_{k=1}^h \sum_{p \in \mathcal{S}_{k-1}(v)} \frac{f(1)}{\epsilon n} \mathbb{P}(p \text{ is } \theta_n\text{-admissible in } G_n^\epsilon) = \frac{f(1)}{\epsilon n} \sum_{k=0}^{h-1} \mathbb{E}[N_k^{\theta_n}(v)].$$

Lemma 3.11 yields for small  $\bar{\delta} > 0$  and  $\epsilon := \epsilon - \bar{\delta} > 0$ ,

$$\mathbb{P}(N_{\leq h}^{\theta_n}(v, w) \geq 1) \leq \frac{f(1)}{\epsilon n} \sum_{k=0}^{h-1} E_{(s_n(v), \ell)} \left[ |\text{IBP}_k^\epsilon((1 + \bar{\delta})f)| \right].$$

We denote by  $\bar{\rho}$  the spectral radius of the operator  $A$  associated to  $\text{IBP}^\epsilon((1 + \bar{\delta})f)$  and by  $\bar{\varphi}$  the corresponding eigenfunction. Choose a constant  $C$  such that for all sufficiently small  $\bar{\delta}$ ,  $C \geq \max_\phi \bar{\varphi}(\phi) / \min_\phi \bar{\varphi}(\phi)$ . This is possible since the eigenfunctions are continuous in  $\bar{\delta}$  (this can be seen along the lines of Note 3 to Chapter II on pages 568-569 of [27]). Furthermore, by the continuity of the spectral radius with respect to the operator (see Chapter II.5 in [27]) and, since  $\rho_\epsilon(A) > 1$  by assumption,  $\bar{\rho} > 1$  for all small  $\bar{\delta}$ . Hence, for all  $v, w \in V_n^\epsilon$ ,  $v \neq w$ ,

$$\begin{aligned} \mathbb{P}(N_{\leq h}^{\theta_n}(v, w) \geq 1) &\leq \frac{f(1)}{\epsilon n} \sum_{k=0}^{h-1} \frac{1}{\min_\phi \bar{\varphi}(\phi)} E_{(s_n(v), \ell)} \left[ \sum_{\substack{x \in \text{IBP}^\epsilon((1 + \bar{\delta})f) \\ |x|=k}} \bar{\varphi}(\phi) \right] \\ &= \frac{f(1)}{\epsilon n} \sum_{k=0}^{h-1} \bar{\rho}^k \frac{\bar{\varphi}(s_n(v), \ell)}{\min_\phi \bar{\varphi}(\phi)} \\ &\leq \frac{f(1)C}{\epsilon n} \frac{\bar{\rho}^h}{\bar{\rho} - 1} = \frac{f(1)C}{\epsilon(\bar{\rho} - 1)} \exp(h \log \bar{\rho} - \log n). \end{aligned}$$

In particular, for  $\delta > 0$  and  $h_n := (1 - \delta^2) \frac{\log n}{\log \bar{\rho}}$ , we showed that

$$\sup_{v, w \in V_n^\epsilon, v \neq w} \mathbb{P}(N_{\leq h_n}^{\theta_n}(v, w) \geq 1) = o(1).$$

For independent, uniformly chosen vertices  $V_n, W_n$  in  $C_n^\epsilon$ , we have  $V_n \neq W_n$  with high probability. According to (3.10), this implies  $\mathbb{P}(\text{d}_{G_n^\epsilon}(V_n, W_n) \leq h_n) = o(1)$ . Choosing  $\bar{\delta}$  so small that  $\log \bar{\rho} \leq (1 + \delta) \log \rho_\epsilon(A)$ , it follows that, with high probability,

$$\text{d}_{G_n^\epsilon}(V_n, W_n) \geq (1 - \delta^2) \frac{\log n}{\log \bar{\rho}} \geq (1 - \delta) \frac{\log n}{\log \rho_\epsilon(A)}. \quad \square$$

## 4 Approximation by a branching process

In this section, we compare the connected components in the network to the multi-type branching process defined in Section 2.1. We begin by coupling the local neighbourhood of a uniformly chosen vertex to the IBP in Sections 4.1 and 4.2. This local

consideration allows us to draw conclusions about the existence or nonexistence of the giant component from knowledge of the branching process, see Section 4.4. For the analysis of the typical distances in the network, knowing the local neighbourhood is insufficient. We show in Section 4.3 that a slightly larger IBP dominates the network globally in a suitable way.

#### 4.1 Coupling the network to a tree

The proof of the coupling follows the lines of [15] for the undamaged network, but unfortunately we cannot use their results directly as the coupling in [15] makes extensive use of vertices which are removed in the damaged network. Note however that the removal of the old vertices significantly reduces the risk of cycles in the local neighbourhood of a vertex and, therefore, the coupling here will be successful for much longer than the coupling in [15].

In the first step, we couple the local neighbourhood of a vertex  $v_0$  in  $G_n^\epsilon$  to a labelled tree  $T_n^\epsilon(v_0)$ , thus ruling out cycles in that subgraph. In Section 4.2 we then study the asymptotics of the offspring distributions to arrive at the IBP.

Every vertex  $v$  in the labelled tree  $T_n^\epsilon(v_0)$  is equipped with a ‘tag’ in  $V_n^\epsilon$  and a ‘mark’  $\alpha \in V_n^\epsilon \cup \{\ell\}$ . The tag indicates which vertex in the network is approximated by  $v$ . We use the same notation for vertex and tag to emphasize the similarity between the tree and the network. The mark  $\alpha$  carries information about the tag of the parent  $w$  of  $v$  in the tree. In the spirit of Section 1.6,  $v$  has mark  $\alpha = \ell$  if its parent has a smaller tag, i.e.  $w < v$ , and we say that the parent of  $v$  is on its left. In contrast, if  $w > v$  we say that the parent is on its right. It turns out that here it is beneficial to record the exact tag of  $w$  instead of only the relative position and we choose  $\alpha = w$ . Hence, a typical label is of the form  $(v, \alpha)$ .

To construct the coupling, we run an exploration process on the connected component of  $v_0$ . The offspring distribution of a vertex  $v$  in the tree is chosen to be the same as the distribution of direct neighbours of  $v$  in  $G_n^\epsilon$  when only the vertex  $w$  is known as whose direct neighbour  $v$  is found in the exploration. That vertex  $w$  determines the mark of  $v$ . The need of this information to identify the offspring distribution is the reason why vertices in  $T_n^\epsilon(v_0)$  are equipped with marks, whereas vertices in  $G_n^\epsilon(v_0)$  are not. Note the similarity to the comparison between network and IBP sketched in Section 1.6.

Formally, for  $v_0 \in V_n^\epsilon$ , let  $T_n^\epsilon(v_0)$  be the random tree with root  $v_0$  of label  $(v_0, \ell)$  constructed as follows: Every vertex  $v$  produces independently offspring to the left, i.e., with tag  $u \in \{\lfloor \epsilon n \rfloor + 1, \dots, v - 1\}$  with probability

$$\mathbb{P}(v \text{ has a descendant with tag } u) = \mathbb{P}(\Delta \mathcal{Z}[u, v - 1] = 1).$$

All offspring on the left are of mark  $v$ . Moreover, independently,  $v$  produces descendants to its right (i.e. with tag at least  $v+1$ ). Since the parent of these descendant is on the left, they are of mark  $\ell$ . The distribution of the cumulative sum<sup>2</sup> of the sequence of relative positions of the right descendants depends on the mark of  $v$ . When  $v$  is of mark  $\alpha = \ell$ , then the cumulative sum is distributed according to the law of  $(\mathcal{Z}[v, u]: v + 1 \leq u \leq n)$ . When  $v$  is of mark  $\alpha = w \in V_n^\epsilon$ ,  $w > v$ , then the cumulative sum follows the same distribution as  $(\mathcal{Z}[v, u] - \mathbb{1}_{[w, \infty)}(u): v + 1 \leq u \leq n)$  conditioned on  $\Delta \mathcal{Z}[v, w - 1] = 1$ . The percolated version  $T_{n,p}^\epsilon(v_0)$  is obtained from  $T_n^\epsilon(v_0)$  by deleting every particle in  $T_n^\epsilon(v_0)$  together with its line of descent with probability  $1 - p$ , independently for all particles. In particular, with probability  $1 - p$  the root  $v_0$  is deleted and  $T_{n,p}^\epsilon(v_0)$  is empty.

We write  $C_{n,p}^\epsilon(v_0)$  for the connected component in  $G_n^\epsilon(p)$  containing vertex  $v_0$ .

<sup>2</sup>For a sequence  $(x_j: j = 1, \dots, n)$  the cumulative sum is given by  $(\sum_{i=1}^j x_i: j = 1, \dots, n)$

**Proposition 4.1.** *Suppose  $(c_n: n \in \mathbb{N})$  is a sequence of positive integers that satisfies  $\lim_{n \rightarrow \infty} c_n^2/n = 0$ . Then there exists a coupling of a uniformly chosen vertex  $V_n$  in  $V_n^\epsilon$ , graph  $G_n^\epsilon(p)$  and tree  $T_{n,p}^\epsilon(V_n)$  such that*

$$|C_{n,p}^\epsilon(V_n)| \wedge c_n = |T_{n,p}^\epsilon(V_n)| \wedge c_n \quad \text{with high probability.}$$

To prove Proposition 4.1, we define an exploration process which we then use to inductively collect information about the tree and the network on the same probability space. We show that the two discovered graphs agree until a stopping time, which is with high probability larger than  $c_n$ . After that time, the undiscovered part of the tree and the network can be generated independently of each other. This result offers a detailed description of the local neighbourhood of a vertex, and is much stronger than the equal cardinality stated in Proposition 4.1.

We begin by specifying the exploration process that is used to explore the connected component of a vertex  $v_0$  in a labelled graph  $G$ , like  $C_{n,p}^\epsilon(v_0)$  or  $T_{n,p}^\epsilon(v_0)$ . We distinguish three categories of vertices:

- *veiled vertices*: vertices for which we have not yet found a connection to the cluster of  $v_0$
- *active vertices*: vertices for which we already know that they belong to the cluster of  $v_0$  but for which we have not yet explored all its immediate neighbours
- *dead vertices*: vertices which belong to the cluster of  $v_0$  and for which all immediate neighbours have been explored

At the beginning of the exploration only  $v_0$  is active and all other vertices are veiled. In the first exploration step we explore all immediate neighbours of  $v_0$ , declare  $v_0$  as dead and all its immediate neighbours as active. The other vertices remain veiled. We now continue from the active vertex  $v$  with the smallest tag and explore all its immediate neighbours apart from  $v_0$  from where we just came. The exploration is continued until there are no active vertices left.

We couple the exploration processes of the network and the tree started with  $v_0 \in V_n^\epsilon$  up to a stopping time  $T$ , such that up to time  $T$  both explored subgraphs (without the marks) coincide. In particular, the explored part of  $C_{n,p}^\epsilon(v_0)$  is a tree and every tag has been used at most once by the active or dead vertices in  $T_{n,p}^\epsilon(v_0)$ . The event that at least one of these properties fails is called  $E$ . We also stop the exploration, when either the number of dead and active vertices exceeds  $c_n$  or when there are no active vertices left. In this case we say that the *coupling is successful*. If we have to stop as a consequence of  $E$ , we say that the *coupling fails*.

**Lemma 4.2.** *Suppose that  $p \in (0, 1]$  and  $(c_n: n \in \mathbb{N})$  satisfies  $\lim_{n \rightarrow \infty} c_n^2/n = 0$ . Then*

$$\lim_{n \rightarrow \infty} \sup_{v_0 \in V_n^\epsilon} \mathbb{P}(\text{the coupling of } C_{n,p}^\epsilon(v_0) \text{ and } T_{n,p}^\epsilon(v_0) \text{ fails}) = 0.$$

In the sequel, we will label some key constants by the lemma in which they appear first. In the proof of Lemma 4.2 we make use of the following result.

**Lemma 4.3** (Adaption of Lemma 2.12 in [15]). *Let  $(c_n: n \in \mathbb{N})$  satisfy  $\lim_{n \rightarrow \infty} c_n/n = 0$ . Then there exists a constant  $C_{4.3} > 0$  such that for all sufficiently large  $n$ , for all disjoint sets  $\mathcal{I}_0, \mathcal{I}_1 \subseteq V_n^\epsilon$  with  $|\mathcal{I}_0| \leq c_n$  and  $|\mathcal{I}_1| \leq 1$ , and for all  $u, v \in V_n^\epsilon$ ,*

$$\begin{aligned} \mathbb{P}(\Delta \mathcal{Z}[v, u] = 1 \mid \Delta \mathcal{Z}[v, i] = 1 \text{ for } i \in \mathcal{I}_1, \Delta \mathcal{Z}[v, i] = 0 \text{ for } i \in \mathcal{I}_0) \\ \leq C_{4.3} \mathbb{P}(\Delta \mathcal{Z}[v, u] = 1 \mid \Delta \mathcal{Z}[v, i] = 1 \text{ for } i \in \mathcal{I}_1). \end{aligned}$$

*Proof.* We have

$$\begin{aligned} \mathbb{P}(\Delta\mathcal{Z}[v, u] = 1 \mid \Delta\mathcal{Z}[v, i] = 1 \text{ for } i \in \mathcal{I}_1, \Delta\mathcal{Z}[v, i] = 0 \text{ for } i \in \mathcal{I}_0) \\ \leq \frac{\mathbb{P}(\Delta\mathcal{Z}[v, u] = 1 \mid \Delta\mathcal{Z}[v, i] = 1 \text{ for } i \in \mathcal{I}_1)}{\mathbb{P}(\Delta\mathcal{Z}[v, i] = 0 \text{ for } i \in \mathcal{I}_0 \mid \Delta\mathcal{Z}[v, i] = 1 \text{ for } i \in \mathcal{I}_1)}. \end{aligned}$$

With  $n$  so large that  $\lfloor \epsilon n \rfloor \geq \hat{n}$ , Lemma 3.9 and (3.8) imply that

$$\begin{aligned} \mathbb{P}(\Delta\mathcal{Z}[v, i] = 0 \text{ for } i \in \mathcal{I}_0 \mid \Delta\mathcal{Z}[v, i] = 1 \text{ for } i \in \mathcal{I}_1) &\geq \mathbb{P}^1(\Delta\mathcal{Z}[v, i] = 0 \text{ for } i \in \mathcal{I}_0) \\ &\geq \prod_{i \in \mathcal{I}_0} \mathbb{P}^1(\Delta\mathcal{Z}[v, i] = 0) \geq \prod_{i \in \mathcal{I}_0} \left(1 - \frac{f(1)}{(v-1)^\gamma i^{1-\gamma}}\right) \geq \left(1 - \frac{f(1)}{\epsilon n}\right)^{c_n}. \end{aligned}$$

Since  $c_n/n$  tends to zero as  $n \rightarrow \infty$ , the right-hand side converges to one.  $\square$

*Proof of Lemma 4.2.* We assume that  $n$  is so large that  $\lfloor \epsilon n \rfloor \geq \hat{n}$ . To distinguish the exploration processes, we use the term *descendant* for a child in the labelled tree and the term *neighbour* in the context of  $G_n^\epsilon(p)$ . The  $\sigma$ -algebra generated by the exploration until the completion of step  $k$  is denoted  $\mathcal{F}_k$ .

Since the probability of removing  $v_0$  is the same in  $C_{n,p}^\epsilon(v_0)$  and  $T_{n,p}^\epsilon(v_0)$ , this event can be perfectly coupled. If  $v_0$  is not removed, then we explore the immediate neighbours of  $v_0$  in  $G_n^\epsilon(p)$  and the children of the root  $v_0$  in the tree. Again these families are identically distributed and can be perfectly coupled.

Now suppose that we successfully completed exploration step  $k$  and are about to start the next step from vertex  $v$ . At this stage every vertex in the tree can be uniquely referred to by its tag and the subgraphs coincide. Denoting by  $\mathfrak{a}$  and  $\mathfrak{d}$  the set of active and dead vertices, respectively, we have  $\mathfrak{a} \neq \emptyset$  and  $|\mathfrak{a} \cup \mathfrak{d}| < c_n$ . We continue by exploring the left descendants and neighbours of  $v$ . Since we always explore the leftmost active vertex, we cannot encounter a dead or active neighbour in this step. However, in the tree  $T_{n,p}^\epsilon(v_0)$  we may find a dead left descendant (i.e. an offspring whose tag agrees with the tag of a dead particle); we call this event  $\text{Ia}$ . On  $\text{Ia}$ , the vertices in the explored part of  $T_{n,p}^\epsilon(v_0)$  are no longer uniquely identifiable by their tag and we stop. We have

$$\mathbb{P}(\text{Ia} \mid \mathcal{F}_k) = \mathbb{P}(\exists d \in \mathfrak{d} : d \text{ left descendant of } v \mid \mathcal{F}_k) \leq \sum_{d \in \mathfrak{d}} \mathbb{P}(\Delta\mathcal{Z}[d, v-1] = 1) \leq c_n \frac{f(0)}{\epsilon n}.$$

In the first inequality, we used subadditivity, the definition of  $T_{n,p}^\epsilon(v_0)$  and omitted the event that offspring of  $v$  are removed by percolation. Hence,  $\mathbb{P}(\text{Ia}) = O(c_n/n)$ . In the exploration to the left in the tree, we immediately check if a found left descendant has a right descendant which is dead. We denote this event by  $\text{Ib}$  and stop the exploration as soon as it occurs. The reason is that in the network this event could not happen since we always explore the leftmost active vertex. Therefore, the distribution of left neighbours agrees with the distribution of the left descendants conditioned on having no dead right descendants and we can couple both explorations such that they agree in this case. To estimate the probability of the adverse event  $\text{Ib}$ , we use the definition of  $T_{n,p}^\epsilon(v)$  to obtain

$$\begin{aligned} \mathbb{P}(\text{Ib} \mid \mathcal{F}_k) &\leq \mathbb{P}(\exists u \in \mathfrak{d}^c, d \in \mathfrak{d} : u \text{ left descendant of } v, d \text{ right descendant of } u \mid \mathcal{F}_k) \\ &\leq \sum_{u \in \mathfrak{d}^c} \sum_{d \in \mathfrak{d}} \mathbb{P}(\Delta\mathcal{Z}[u, v-1] = 1) \mathbb{P}(\Delta\mathcal{Z}[u, d-1] = 1 \mid \Delta\mathcal{Z}[u, v-1] = 1). \end{aligned}$$

By definition of the exploration process, there are at most  $c_n$  dead vertices. Therefore, Lemma 3.9 and (3.8) yield

$$\mathbb{P}(\text{Ib} \mid \mathcal{F}_k) \leq c_n \sum_{u \in \mathfrak{d}^c, u \leq v-1} \frac{f(0)}{(u-1)^\gamma (v-1)^{1-\gamma}} \frac{f(1)}{\epsilon n} \leq c_n \frac{f(0)f(1)}{\epsilon n} \frac{1}{(v-1)^{1-\gamma}} \sum_{u=1}^{v-1} u^{-\gamma},$$

which implies in particular that  $\mathbb{P}(\text{Ib}) = O(c_n/n)$ .

We turn to the exploration of right descendants, resp. neighbours. When vertex  $v$  is of mark  $\alpha \neq \ell$ , then we already know that  $v$  has no right descendants, resp. neighbours, in  $\mathfrak{d}$  since we checked this when  $v$  was discovered. We denote the event that a right descendant, resp. neighbour, is active by  $\text{Iir}$  and stop the exploration as soon as this event occurs because the tags in  $\mathbb{T}_{n,p}^\epsilon(v_0)$  are no longer unique, resp. we found a cycle in  $\mathbb{C}_{n,p}^\epsilon(v_0)$ . According to Lemma 4.3 and (3.8),

$$\begin{aligned} \mathbb{P}(\text{Iir} \mid \mathcal{F}_k) &\leq \mathbb{P}(\exists a \in \mathfrak{a}: \Delta \mathcal{Z}[v, a-1] = 1 \mid \Delta \mathcal{Z}[v, \alpha-1] = 1, \Delta \mathcal{Z}[v, d-1] = 0 \text{ for } d \in \mathfrak{d} \setminus \{\alpha\}, \mathcal{F}_k) \\ &\leq C_{4.3} \sum_{a \in \mathfrak{a}} \mathbb{P}^1(\Delta \mathcal{Z}[v, a-1] = 1) \leq C_{4.3} c_n \frac{f(1)}{\epsilon n}. \end{aligned}$$

Thus,  $\mathbb{P}(\text{Iir}) = O(c_n/n)$ . Conditional on the event that there are no active vertices in the set of right descendants, resp. neighbours, the offspring distribution in tree and network agree and can therefore be perfectly coupled. When the vertex  $v$  is of mark  $\alpha = \ell$ , then we have not gained any information about its right descendants, yet. The event that there is a dead or active vertex in the right descendants is denoted by  $\text{IIla}$ . We stop when this event occurs and use (3.8) to estimate

$$\begin{aligned} \mathbb{P}(\text{IIla} \mid \mathcal{F}_k) &= \mathbb{P}(\exists a \in \mathfrak{a} \cup \mathfrak{d}: a \text{ right descendant of } v \mid \mathcal{F}_k) \\ &\leq \sum_{a \in \mathfrak{a} \cup \mathfrak{d}} \mathbb{P}(\Delta \mathcal{Z}[v, a-1] = 1) \leq c_n \frac{f(0)}{\epsilon n}. \end{aligned}$$

Thus,  $\mathbb{P}(\text{IIla}) = O(c_n/n)$ . In  $\mathbb{C}_{n,p}^\epsilon(v_0)$ , we know that  $v$  has no dead right neighbours as this would have stopped the exploration in the moment when  $v$  became active. The event that there are active vertices in the set of right neighbours is denoted by  $\text{IIlb}$  and we stop as soon as it occurs since a cycle is created. Using again (3.8), we find

$$\begin{aligned} \mathbb{P}(\text{IIlb} \mid \mathcal{F}_k) &= \mathbb{P}(\exists a \in \mathfrak{a}: \Delta \mathcal{Z}[v, a-1] = 1 \mid \Delta \mathcal{Z}[v, d-1] = 0 \text{ for } d \in \mathfrak{d}, \mathcal{F}_k) \\ &\leq \sum_{a \in \mathfrak{a}} \mathbb{P}(\Delta \mathcal{Z}[v, a-1] = 1) \leq c_n \frac{f(0)}{\epsilon n}. \end{aligned}$$

As in case  $\alpha \neq \ell$ , the explorations can be perfectly coupled when the adverse events do not occur. We showed that in every step the coupling fails with a probability bounded by  $O(c_n/n)$ . As there are at most  $c_n$  exploration steps until we end the coupling successfully, the probability of failure is  $O(c_n^2/n) = o(1)$ . In other words, the coupling succeeds with high probability.  $\square$

**Proof of Proposition 4.1.** First, consider the statement for a fixed vertex  $v_0$ . When the coupling is successful and ends because at least  $c_n$  vertices were explored, then  $|\mathbb{C}_{n,p}^\epsilon(v_0)| \geq c_n$  and  $|\mathbb{T}_{n,p}^\epsilon(v_0)| \geq c_n$ . If the coupling is successful and ends because there are no active vertices left, then  $|\mathbb{C}_{n,p}^\epsilon(v_0)| = |\mathbb{T}_{n,p}^\epsilon(v_0)|$  since the subgraphs coincide. Since the coupling is successful with high probability by Lemma 4.2,  $|\mathbb{C}_{n,p}^\epsilon(v_0)| \wedge c_n = |\mathbb{T}_{n,p}^\epsilon(v_0)| \wedge c_n$  with high probability. As Lemma 4.2 shows the success of the coupling uniformly in the start vertex, the randomization of the vertex  $v_0$  to a uniformly chosen vertex  $V_n \in \mathbb{V}_n^\epsilon$  is now straightforward.  $\square$

## 4.2 Coupling the tree to the IBP

Coupling the neighbourhood of a vertex to a labelled tree provides a great simplification of the problem since many dependencies are eliminated. However, the offspring

distribution in the tree  $\mathbb{T}_{n,p}^\epsilon(V_n)$  is still complicated and depends on  $n$ . Since we are mainly interested in the asymptotic size of the giant component, we now couple the tree to the IBP, which does not depend on  $n$  and is much easier to analyse. We denote by  $|\mathcal{X}^\epsilon(p)|$  the total progeny of the IBP. Recall the definition of  $S^\epsilon$  from (2.2).

**Proposition 4.4.** *Let  $p \in (0, 1]$  and  $(c_n: n \in \mathbb{N})$  be a sequence of positive integers with  $\lim_{n \rightarrow \infty} c_n^3/n = 0$ . Then there exists a coupling of a uniformly chosen vertex  $V_n$  in  $\mathbb{V}_n^\epsilon$ , the graph  $G_n^\epsilon(p)$  and the percolated IBP started with a particle of mark  $\ell$  and location  $S^\epsilon$  such that, with high probability,*

$$|C_{n,p}^\epsilon(V_n)| \wedge c_n = |\mathcal{X}^\epsilon(p)| \wedge c_n.$$

*Proof of Proposition 4.4.* Throughout the proof, suppose that  $n$  is so large that  $\lfloor \epsilon n \rfloor \geq \hat{n}$ . Instead of coupling the IBP directly to the network, we couple a projected version of the IBP to the tree  $\mathbb{T}_{n,p}^\epsilon(V_n)$ . As long as the number of particles is preserved under the projection, this is sufficient according to Proposition 4.1. To describe the projection, we define  $\pi_n^\epsilon: [\log \epsilon, 0] \rightarrow \mathbb{V}_n^\epsilon$  by

$$\pi_n^\epsilon(\lambda) = v \quad \Leftrightarrow \quad s_n(v-1) < \lambda \leq s_n(v), \quad (4.1)$$

where  $s_n(v) = -\sum_{j=v}^{n-1} \frac{1}{j}$ . Since  $s_n(\lfloor \epsilon n \rfloor) < \log(\lfloor \epsilon n \rfloor/n) \leq \log \epsilon$ , every location in  $[\log \epsilon, 0]$  can be uniquely identified with a tag in  $\mathbb{V}_n^\epsilon$  by the map  $\pi_n^\epsilon$ . The projected IBP is again a labelled tree: The genealogical tree of the IBP with its marks is preserved, the location of a particle  $x$  is replaced by the tag  $\pi_n^\epsilon(\lambda(x))$ . If  $s_n(\lfloor \epsilon n \rfloor + 1) < \log \epsilon$ , then no particles of the IBP are projected onto  $\lfloor \epsilon n \rfloor + 1$ . Moreover, while for  $v \geq \lfloor \epsilon n \rfloor + 3$  an interval of length  $1/(v-1)$  is projected onto  $v$ , for  $\lfloor \epsilon n \rfloor + 2$  only an interval of length at most  $s_n(\lfloor \epsilon n \rfloor + 2) - \log \epsilon$  is used. This length is positive but may be smaller than  $1/(\lfloor \epsilon n \rfloor + 1)$ . As a consequence, the projected IBP can have unusually few particles at  $\lfloor \epsilon n \rfloor + 1$  and  $\lfloor \epsilon n \rfloor + 2$  and we treat these two tags separately.

The exploration of the two trees follows the same procedure as the exploration described in Section 4.1 and we declare the coupling successful and stop as soon as either there are no active vertices left or the number of active and dead vertices exceeds  $c_n$ . Since both objects are trees, as long as the labels for the starting vertices agree, any failure of the coupling comes from a failure in the coupling of the offspring distributions. For simplicity, we consider only the case  $p = 1$ . The generalization to  $p \in (0, 1]$  is straightforward.

We first show that the labels of the starting vertices can be coupled with high probability. To this end, note that the distribution of  $S^\epsilon$  is chosen such that  $\exp(S^\epsilon)$  is uniformly distributed on  $[\epsilon, 1]$ . Since  $\log \epsilon \leq s_n(\lfloor \epsilon n \rfloor + 2) \leq s_n(v-1) \leq s_n(v) \leq 0$ , for  $v \geq \lfloor \epsilon n \rfloor + 3$ , we obtain

$$\begin{aligned} \mathbb{P}(\pi_n^\epsilon(S^\epsilon) = v) &= \mathbb{P}(e^{s_n(v-1)} < e^{S^\epsilon} \leq e^{s_n(v)}) \\ &= \frac{1}{1-\epsilon} (e^{s_n(v)} - e^{s_n(v-1)}) = \frac{1}{1-\epsilon} e^{s_n(v-1)} (e^{1/(v-1)} - 1). \end{aligned}$$

The right-hand side is in the interval  $[\frac{1}{1-\epsilon}(\frac{1}{n} - \frac{2}{vn}), \frac{1}{1-\epsilon}(\frac{1}{n} + \frac{2}{vn})]$ . Moreover, the probability that  $V_n$  or  $\pi_n^\epsilon(S^\epsilon)$  is in  $\{\lfloor \epsilon n \rfloor + 1, \lfloor \epsilon n \rfloor + 2\}$  is of order  $O(1/n)$ . Hence,  $V_n$  and  $S^\epsilon$  can be coupled such that

$$\mathbb{P}(V_n \neq \pi_n^\epsilon(S^\epsilon)) \leq \sum_{v=\lfloor \epsilon n \rfloor+3}^n \left| \mathbb{P}(\pi_n^\epsilon(S^\epsilon) = v) - \frac{1}{n-\lfloor \epsilon n \rfloor} \right| + O\left(\frac{1}{n}\right) = O\left(\frac{\log n}{n}\right).$$

In the next step we study the offspring distributions of a particle  $x$  in the IBP with label  $(\lambda, \alpha)$  and  $\pi_n^\epsilon(\lambda) = v$ . We start with the offspring to the left. Let  $u \in \{\lfloor \epsilon n \rfloor + 1, \dots, v\}$ .

By definition of the IBP, particle  $x$  produces a Poissonian number of projected offspring with tag  $u$  with parameter

$$\int_{(s_n(u-1)-\lambda) \vee (\log \epsilon - \lambda)}^{(s_n(u)-\lambda) \wedge 0} \beta e^{(1-\gamma)t} dt.$$

A vertex with tag  $v$  in  $\mathbb{T}_n^\epsilon(V_n)$  produces a Bernoulli distributed number of descendants with tag  $u$  with success probability  $\mathbb{P}(\Delta \mathcal{Z}[u, v-1] = 1)$  when  $u < v$ , and with success probability zero when  $u = v$ . It is proved in Lemma 6.3 of [15] that for  $u \geq \lfloor \epsilon n \rfloor + 3$  the Poisson distributions can be coupled to the Bernoulli distribution such that they disagree with a probability bounded by a constant multiple of  $v^{\gamma-1} u^{-(\gamma+1)}$  for  $u < v$  and  $1/v$  for  $u = v$ . For  $u \in \{\lfloor \epsilon n \rfloor + 1, \lfloor \epsilon n \rfloor + 2\}$  a similar estimate shows that the probability can be bounded by a constant multiple of  $1/(\epsilon n)$ . Since the number of descendants with tag in  $\{\lfloor \epsilon n \rfloor + 1, \dots, v\}$  form an independent sequence of random variables, we can apply the coupling sequentially for each location and obtain a coupling of the  $\pi_n^\epsilon$ -projected left descendants in the IBP and the left descendants in  $\mathbb{T}_n^\epsilon(V_n)$ . The failure probability of this coupling can be estimated by

$$\begin{aligned} \mathbb{P}(\text{left descendants of } v \text{ disagree}) &\leq \frac{3C}{\epsilon n} + \frac{C'}{v^{1-\gamma}} \sum_{u=\lfloor \epsilon n \rfloor + 3}^{v-1} \frac{1}{u^{\gamma+1}} \\ &\leq \frac{3C}{\epsilon n} + \frac{C'}{\epsilon n} \log \left( \frac{v-1}{\lfloor \epsilon n \rfloor} \right) \leq \frac{C''}{n}, \end{aligned}$$

where  $C, C', C''$  are suitable positive constants whose value can change from line to line in the sequel. We turn to the offspring on the right. Suppose that particle  $x$  in the IBP has mark  $\alpha = \ell$ . The cumulative sum of  $\pi_n^\epsilon$ -projected right descendants of  $x$  follows the same distribution as  $(Z_{s_n(u)-\lambda} : v \leq u \leq n)$ . The cumulative sum of right descendants of  $v$  in  $\mathbb{T}_n^\epsilon(V_n)$  is distributed according to the law of  $(\mathcal{Z}[v, u] : v \leq u \leq n)$ . The following lemma is taken from [15] and we omit its proof.

**Lemma 4.5** (Lemma 6.2 in [15]). *Fix a level  $H \in \mathbb{N}$ . We can couple the processes  $(Z_{s_n(u)-\lambda} : v \leq u \leq n)$  and  $(\mathcal{Z}[v, u] : v \leq u \leq n)$  such that for the coupled processes  $(\mathcal{Y}_u^{(1)} : v \leq u \leq n)$  and  $(\mathcal{Y}_u^{(2)} : v \leq u \leq n)$*

$$\mathbb{P}(\mathcal{Y}_u^{(1)} \neq \mathcal{Y}_u^{(2)} \text{ for some } u \leq \sigma_H) \leq C_{4.5} \frac{f(H)^2}{v-1},$$

for some constant  $C_{4.5} > 0$  and  $\sigma_H$  the first time one of the processes reaches or exceeds  $H$ .

In the coupling between the tree  $\mathbb{T}_n^\epsilon(V_n)$  and the projected IBP we consider at most  $c_n$  right descendants. Hence, Lemma 4.5 implies that the distributions can be coupled such that

$$\mathbb{P}(\text{right descendants of } v \text{ disagree}) \leq C_{4.5} \frac{f(c_n)^2}{v-1} \leq C \frac{c_n^2}{\epsilon n},$$

for some  $C > 0$ . When  $\alpha = r$ , then cumulative sum of  $\pi_n^\epsilon$ -projected right descendants of  $x$  follows the same distribution as  $(\hat{Z}_{s_n(u)-\lambda} - 1 : v \leq u \leq n)$ . The cumulative sum of right descendants of a vertex  $v$  with mark  $w \in V_n^\epsilon$ ,  $w > v$ , in  $\mathbb{T}_n^\epsilon(V_n)$  is distributed according to  $(\mathcal{Z}[v, u] - \mathbb{1}_{[w, \infty)}(u) : v \leq u \leq n)$  conditioned on  $\Delta \mathcal{Z}[v, w-1] = 1$ . We can couple these two distributions. Again the proof of the following lemma is up to minor changes given in [15] and therefore omitted.

**Lemma 4.6** (Lemma 6.6 in [15]). *Fix a level  $H \in \mathbb{N}$ . We can couple the processes  $(\hat{Z}_{s_n(u)-\lambda} - 1 : v \leq u \leq n)$  and  $(\mathcal{Z}[v, u] - \mathbb{1}_{[w, \infty)}(u) : v \leq u \leq n)$  conditioned on  $\Delta \mathcal{Z}[v, w - 1] = 1$  such that for the coupled processes  $(\mathcal{Y}_u^{(1)} : v \leq u \leq n)$  and  $(\mathcal{Y}_u^{(2)} : v \leq u \leq n)$*

$$\mathbb{P}(\mathcal{Y}_u^{(1)} \neq \mathcal{Y}_u^{(2)} \text{ for some } u \leq \sigma_H) \leq C_{4.6} \frac{f(H)^2}{v-1},$$

for some constant  $C_{4.6} > 0$  and  $\sigma_H$  the first time one of the processes reaches or exceeds  $H$ .

As we explore at most  $c_n$  vertices during the exploration, Lemma 4.6 implies that we can couple the offspring distribution to the right such that there is a constant  $C > 0$  with

$$\mathbb{P}(\text{right descendants of } v \text{ disagree}) \leq C_{4.6} \frac{f(c_n)^2}{v-1} \leq C \frac{c_n^2}{\epsilon n}.$$

Since we explore at most  $c_n$  vertices in total, the probability that the coupling fails can be bounded by a constant multiple of  $c_n/n + c_n^3/n$ , which converges to zero. Thus, the two explorations can be successfully coupled with high probability and, as in the proof of Proposition 4.1, the claim follows.  $\square$

### 4.3 Dominating the network by a branching process

Like in the coupling, we begin with a comparison to a tree: For  $\theta \in \mathbb{N}$  and  $v_0 \in V_n^\epsilon$ , let  $\mathbb{T}_n^{\epsilon, \theta}(v_0)$  be the subtree of  $\mathbb{T}_n^\epsilon(v_0)$ , where every particle can have at most  $\theta$  offspring to the right. That is, for a particle with tag  $v$  and mark  $\alpha = \ell$  the cumulative sum of the offspring to the right is distributed according to the law of  $(\mathcal{Z}[v, u] \wedge \theta : v+1 \leq u \leq n)$ . When  $v$  is of mark  $\alpha = w \in V_n^\epsilon$ ,  $w > v$ , then the cumulative sum follows the same distribution as  $((\mathcal{Z}[v, u] - \mathbb{1}_{[w, \infty)}(u)) \wedge \theta : v+1 \leq u \leq n)$  conditioned on  $\Delta \mathcal{Z}[v, w-1] = 1$ . Recall from Section 3.2 that  $N_h^\theta(v_0)$  denotes the number of  $\theta$ -admissible paths of length  $h$  in  $G_n^\epsilon$  with starting-point  $v_0$ .

**Lemma 4.7.** *For all  $\theta, h, n \in \mathbb{N}$ ,  $v_0 \in V_n^\epsilon$ ,*

$$\mathbb{E}[N_h^\theta(v_0)] \leq \mathbb{E}\left[|\{\text{particles in generation } h \text{ of } \mathbb{T}_n^{\epsilon, \theta}(v_0)\}|\right].$$

*Proof.* Let  $\mathbf{p} = (v_0, \dots, v_h) \in \mathcal{S}_h(v_0)$ . Using the notation and set-up from the proof of Theorem 1.3, and the definition of the tree  $\mathbb{T}_n^{\epsilon, \theta}(v_0)$ , one easily checks that in cases A, B, C, E and F of Figure 4 on page 31,  $\mathbb{P}(\mathcal{E}_h | \cap_{i=1}^{h-1} \mathcal{E}_i)$  agrees with the probability that in tree  $\mathbb{T}_n^{\epsilon, \theta}(v_0)$  a vertex with tag  $v_{h-1}$  gives birth to a particle of tag  $v_h$  given that its parent has tag  $v_{h-2}$ . In case D of Figure 4, the tree  $\mathbb{T}_n^{\epsilon, \theta}$  is allowed to have one more offspring on its right, because the edge  $\{v_{h-2}, v_{h-1}\}$  is not accounted for. Hence  $\mathbb{P}(\mathcal{E}_h | \cap_{i=1}^{h-1} \mathcal{E}_i)$  is bounded from above by the probability for the event in the tree. We obtain

$$\mathbb{E}[N_h^\theta(v_0)] = \sum_{\mathbf{p} \in \mathcal{S}_h(v_0)} \mathbb{P}(\mathbf{p} \text{ is } \theta\text{-admissible in } G_n^\epsilon) \leq \sum_{\mathbf{p} \in \mathcal{S}_h(v_0)} \mathbb{P}(\mathbf{p} \text{ present in } \mathbb{T}_n^{\epsilon, \theta}(v_0)).$$

Particles in generation  $h$  of  $\mathbb{T}_n^{\epsilon, \theta}(v_0)$ , who have two ancestors with the same tag, are not represented in the sum on the right-hand side. Adding these, we obtain the result.  $\square$

*Proof of Lemma 3.11.* By Lemma 4.7, it suffices to dominate  $\mathbb{T}_n^{\epsilon, \theta_n}(v_0)$  by the  $\text{IBP}^\epsilon((1 + \delta)f)$  started in  $s_n(v_0)$ , or, as in the proof of Proposition 4.4, by the  $\pi_n^\epsilon$ -projected IBP defined by (4.1). Since both processes are trees starting with the same type of particle, it suffices to compare the offspring distributions. All particles in  $\mathbb{T}_n^{\epsilon, \theta_n}(v_0)$  have a tag  $v > \lfloor \epsilon n \rfloor$ , but the projected IBP can have offspring with tag  $v \in \{\lfloor \epsilon n \rfloor + 1, \dots, \lfloor \epsilon n \rfloor\}$ . Hence, these offspring are ignored in the following, giving us a lower bound on the



projected IBP. We assume that  $n$  is so large, that  $n \geq \hat{n}$  and  $s_n(\lfloor \epsilon n \rfloor + 1) \geq \log \epsilon$ . Let  $x$  be a particle in the IBP of type  $(\lambda, \alpha)$  with  $\pi_n^\epsilon(\lambda) = v$ . We begin with the offspring to the left, i.e. tag  $u \in \{\lfloor \epsilon n \rfloor + 1, \dots, v\}$ . A particle in  $\mathbb{T}_n^{\epsilon, \theta_n}(v_0)$  with tag  $v$  cannot produce particles in  $u = v$ , therefore, the IBP clearly dominates. For  $u < v$ , using (3.8), the probability that a particle with tag  $u$  is a child of  $x$ , is  $\mathbb{P}(\Delta \mathcal{Z}[u, v-1] = 1) \leq \beta(u-1)^{-\gamma}(v-1)^{-(1-\gamma)}$ . Writing  $\bar{f}(k) = (1+\delta)f(k) = \bar{\gamma}k + \bar{\beta}$ , for  $k \in \mathbb{N}_0$ , the number of particles with tag  $u$  produced by  $x$  in the projected IBP follows a Poisson distribution with parameter

$$\int_{s_n(u-1)-\lambda}^{s_n(u)-\lambda} \bar{\beta} e^{(1-\bar{\gamma})t} dt \leq \frac{\bar{\beta}}{u-1} e^{-(1-\bar{\gamma}) \sum_{k=u-1}^{v-1} \frac{1}{k}} \leq \frac{\bar{\beta}}{u-1} \left( \frac{u-2}{v-1} \right)^{1-\gamma},$$

where we used that  $\lambda \leq s_n(v)$  and  $e^y - 1 \geq y$ . For  $\varrho > 0$ ,  $\eta \in [0, 1]$ , the Poisson distribution with parameter  $\varrho$  is dominating the Bernoulli distribution with parameter  $\eta$  if and only if  $e^{-\varrho} \leq 1 - \eta$ . Since  $e^{-y} \leq 1 - y + y^2/2$  for  $y \geq 0$ , it suffices to show that  $\varrho(1 - \varrho/2) \geq \eta$ . In our case,  $\eta = \beta(u-1)^{-\gamma}(v-1)^{-(1-\gamma)}$ ,  $\varrho = \eta(1+\delta)(1 - 1/(u-1))^{1-\gamma}$  and the inequality holds for all large  $n$  and  $u \in V_n^\epsilon$ ,  $u < v$ , since  $\eta$  is a null sequence.

We turn to the right descendants. The pure jump process corresponding to the attachment rule  $\bar{f}$  is denoted by  $\bar{Z}$  and we write  $P^l$  for the distribution of  $\bar{Z}$  when started in  $l$ , that is,  $P^l(\bar{Z}_0 = l) = 1$ . First suppose that  $\alpha = \ell$ . The cumulative sum of  $\pi_n^\epsilon$ -projected right descendants of  $x$  have the distribution of  $(\bar{Z}_{s_n(u)-\lambda} : v \leq u \leq n)$ , where  $\bar{Z}_0 = 0$ . The cumulative sum of right descendants of  $v$  in  $\mathbb{T}_n^{\epsilon, \theta_n}(v_0)$  is distributed according to the law of  $(\mathcal{Z}[v, u] \wedge \theta_n : v \leq u \leq n)$ . We couple these distributions by defining  $((\mathcal{Y}_u^{(1)}, \mathcal{Y}_u^{(2)}) : v \leq u \leq n)$  to be the time-inhomogeneous Markov chain which starts in  $P^0(\bar{Z}_{s_n(v)-\lambda} \in \cdot) \otimes \delta_0$ , has the desired marginals and evolves from state  $(l, k)$  at time  $j$  according to a coupling of  $\bar{Z}_{1/j}$  and  $\mathcal{Z}[j, j+1]$ , which guarantees that  $\bar{Z}_{1/j} \geq \mathcal{Z}[j, j+1]$ , until  $\mathcal{Y}^{(2)}$  reaches state  $\theta_n$ , where  $\mathcal{Y}^{(2)}$  is absorbed. To show that this coupling exists, it suffices to show that

$$e^{-\bar{f}(l)/j} = P^l(\bar{Z}_{1/j} = l) \leq \mathbb{P}^k(\mathcal{Z}[j, j+1] = k) = 1 - f(k)/j \quad \text{for } j \in V_n^\epsilon, k \leq \theta_n, k \leq l.$$

Since  $\bar{f}$  is non decreasing, this inequality follows as above once we show that  $\varrho(1 - \varrho/2) \geq \eta$  with  $\eta = f(k)/j$ ,  $\varrho = \bar{f}(k)/j = \eta(1+\delta)$ . Since  $k \leq \theta_n = o(n)$  and  $j \geq \lfloor \epsilon n \rfloor$ ,  $\eta$  is a null sequence and the claim follows. Hence,  $\mathcal{Y}_j^{(1)} \geq \mathcal{Y}_j^{(2)}$  for all  $j$  and the domination is established.

Now suppose that  $\alpha = r$  and that the location of  $x$  parent is projected onto tag  $w$ . The cumulative sum of  $\pi_n^\epsilon$ -projected right descendants of  $x$  has the distribution of  $(Y_{s_n(u)-\lambda} : v \leq u \leq n)$ , where  $Y$  is a version of  $\bar{Z}$  under measure  $P^1$ . The cumulative sum of right descendants of  $v$  in  $\mathbb{T}_n^{\epsilon, \theta_n}(v_0)$  is distributed according to the law of  $((\mathcal{Z}[v, u] - \mathbb{1}_{[w, \infty)}(u)) \wedge \theta_n : v \leq u \leq n)$  conditioned on  $\Delta \mathcal{Z}[v, w-1] = 1$ . We couple these distributions as in the  $\alpha = \ell$  case, but for times  $j \leq w-2$ , the Markov chain evolves from state  $(l, k)$  according to a coupling of  $Y_{1/j}$  and  $\mathcal{Z}[j, j+1]$  conditioned on  $\Delta \mathcal{Z}[j, w-1] = 1$ , which guarantees that  $Y_{1/j} \geq \mathcal{Z}[j, j+1]$ , until either  $j = w-2$  or  $\mathcal{Y}^{(2)}$  reaches  $\theta_n$  and is absorbed. To show that this coupling exists, it suffices to show that

$$P^{l+1}(\bar{Z}_{1/j} - 1 = l) \leq \mathbb{P}^k(\mathcal{Z}[j, j+1] = k | \mathcal{Z}[j, w-1] = 1) \quad \text{for } j \in V_n^\epsilon, k \leq \theta_n, k \leq l. \quad (4.2)$$

We compute

$$\begin{aligned} \mathbb{P}^k(\mathcal{Z}[j, j+1] = k | \Delta \mathcal{Z}[j, w-1] = 1) &= 1 - \frac{\mathbb{P}^k(\Delta \mathcal{Z}[j, j] = 1, \Delta \mathcal{Z}[j, w-1] = 1)}{\mathbb{P}^k(\Delta \mathcal{Z}[j, w-1] = 1)} \\ &= 1 - \frac{\frac{f(k)}{j} \mathbb{P}^{k+1}(\Delta \mathcal{Z}[j+1, w-1] = 1)}{\mathbb{P}^k(\Delta \mathcal{Z}[j, w-1] = 1)} = 1 - \frac{f(k+1)}{j+\gamma}. \end{aligned}$$

Since  $\bar{f}$  is non-decreasing, (4.2) follows when we show that  $\varrho(1 - \varrho/2) \geq \eta$  with  $\eta = f(k+1)/(j+\gamma)$  and  $\varrho = \bar{f}(k+1)/j = \eta(1+\delta)(1+\gamma/j)$ . Since  $k \leq \theta_n = o(n)$  and  $j \geq \lfloor \epsilon n \rfloor$ ,  $\eta$  is a null sequence and (4.2) is proved. In the transition from generation  $j = w - 1$  to  $j = w$ ,  $\mathcal{Y}^{(2)}$  cannot change its state while  $\mathcal{Y}^{(1)}$  can increase. From generation  $j = w$  onwards, the coupling explained in case  $\alpha = \ell$  is used. Thus, the Markov chain can be constructed such that  $\mathcal{Y}_j^{(1)} \geq \mathcal{Y}_j^{(2)}$  for all  $j$  and the domination is proved.  $\square$

#### 4.4 Proof of Theorem 2.1

Proposition 4.4 implies the following result.

**Corollary 4.8.** *Let  $p \in (0, 1]$  and  $(c_n : n \in \mathbb{N})$  be a sequence with  $\lim_{n \rightarrow \infty} c_n^3/n = 0$  and  $\lim_{n \rightarrow \infty} c_n = \infty$ . Then, as  $n \rightarrow \infty$ ,*

$$\mathbb{E} \left[ \frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{v=\lfloor \epsilon n \rfloor+1}^n \mathbb{1}\{|\mathcal{C}_{n,p}^\epsilon(v)| \geq c_n\} \right] = \mathbb{P}(|\mathcal{C}_{n,p}^\epsilon(V_n)| \geq c_n) \rightarrow P(|\mathcal{X}^\epsilon(p)| = \infty) = \zeta^\epsilon(p).$$

This convergence can be strengthened to convergence in probability.

**Lemma 4.9.** *Let  $p \in (0, 1]$  and  $(c_n : n \in \mathbb{N})$  be a sequence with  $\lim_{n \rightarrow \infty} c_n^3/n = 0$  and  $\lim_{n \rightarrow \infty} c_n = \infty$ . Then*

$$M_{n,p}^\epsilon(c_n) := \frac{1}{(1-\epsilon)n} \sum_{v=\lfloor \epsilon n \rfloor+1}^n \mathbb{1}\{|\mathcal{C}_{n,p}^\epsilon(v)| \geq c_n\} \rightarrow \zeta^\epsilon(p) \quad \text{in probability, as } n \rightarrow \infty.$$

To prove Lemma 4.9, we use a variance estimate for  $M_{n,p}^\epsilon(c_n)$ .

**Lemma 4.10.** *Let  $p \in (0, 1]$  and  $(c_n : n \in \mathbb{N})$  be a positive sequence. There exists a constant  $C > 0$  such that*

$$\text{Var}(M_{n,p}^\epsilon(c_n)) \leq \frac{C}{n} \left( c_n + \frac{c_n^2}{\epsilon n} \right).$$

The proof is almost identical to the proof of Proposition 7.1 in [15]. The necessary changes are similar to the changes made for the proofs of Proposition 4.1 and 4.4. We sketch only the main steps.

*Proof sketch.* Write

$$\begin{aligned} \text{Var} \left( \frac{1}{(1-\epsilon)n} \sum_{v=\lfloor \epsilon n \rfloor+1}^n \mathbb{1}\{|\mathcal{C}_{n,p}^\epsilon(v)| \geq c_n\} \right) &= \frac{1}{(1-\epsilon)^2 n^2} \\ &\cdot \sum_{v,w=\lfloor \epsilon n \rfloor+1}^n \left( \mathbb{P}(|\mathcal{C}_{n,p}^\epsilon(v)| \geq c_n, |\mathcal{C}_{n,p}^\epsilon(w)| \geq c_n) - \mathbb{P}(|\mathcal{C}_{n,p}^\epsilon(v)| \geq c_n) \mathbb{P}(|\mathcal{C}_{n,p}^\epsilon(w)| \geq c_n) \right). \end{aligned} \quad (4.3)$$

To estimate the probability  $\mathbb{P}(|\mathcal{C}_{n,p}^\epsilon(v)| \geq c_n, |\mathcal{C}_{n,p}^\epsilon(w)| \geq c_n)$ , we run two successive explorations in the graph  $G_n^\epsilon(p)$ , the first starting from  $v$  and the second starting from  $w$ . For these explorations, we use the exploration process described below Proposition 4.1 but in every step only neighbours in the set of veiled vertices are explored. The first exploration is terminated as soon as either the number of dead and active vertices exceeds  $c_n$  or there are no active vertices left. The second exploration, additionally, stops when a vertex is found which was already unveiled in the first exploration. We denote

$$\Theta_v := \{\text{the first exploration started in vertex } v \text{ stops because } c_n \text{ vertices are found}\}.$$

Then, for any  $v \in V_n^\epsilon$ ,  $\mathbb{P}(|C_{n,p}^\epsilon(v)| \geq c_n) = \mathbb{P}(\Theta_v)$  and in the proof of Proposition 7.1 of [15] it was shown that there exists a constant  $C' > 0$ , independent of  $v$  and  $n$ , such that

$$\begin{aligned} \sum_{w=\lfloor \epsilon n \rfloor + 1}^n \mathbb{P}(|C_{n,p}^\epsilon(v)| \geq c_n, |C_{n,p}^\epsilon(w)| \geq c_n) \\ \leq \mathbb{P}(\Theta_v) \left( c_n + \sum_{w=\lfloor \epsilon n \rfloor + 1}^n \mathbb{P}(\Theta_w) + C' c_n \mathbb{P}^{c_n}(\Delta \mathcal{Z}[\lfloor \epsilon n \rfloor + 1, \lfloor \epsilon n \rfloor + 1] = 1) \right). \end{aligned} \quad (4.4)$$

Combining (4.3) and (4.4) and using (3.8), we deduce that there exists a constant  $C > 0$  with

$$\text{Var}(M_n^\epsilon(c_n)) \leq \frac{1}{(1-\epsilon)^2 n^2} \sum_{v=\lfloor \epsilon n \rfloor + 1}^n \mathbb{P}(\Theta_v) \left( c_n + C' \frac{c_n f(c_n)}{\epsilon n} \right) \leq \frac{C}{n} \left( c_n + \frac{c_n^2}{\epsilon n} \right). \quad \square$$

**Proof of Lemma 4.9.** Using Chebyshev's inequality, Corollary 4.8 and Lemma 4.10 yield the claim.  $\square$

Lemma 4.9 already implies that the asymptotic relative size of a largest component in the network is bounded from above by  $\zeta^\epsilon(p)$ . To show that the survival probability also constitutes a lower bound, we use the following sprinkling argument.

**Lemma 4.11.** *Let  $\epsilon \in [0, 1)$ ,  $p \in (0, 1]$ ,  $\delta \in (0, f(0))$  and define  $\underline{f}(k) := f(k) - \delta$  for all  $k \in \mathbb{N}_0$ . Denote by  $\underline{C}_{n,p}^\epsilon(v)$  the connected component containing  $v$  in the network  $\underline{G}_n^\epsilon(p)$  constructed with the attachment rule  $\underline{f}$ . Let  $\kappa > 0$  and  $(c_n : n \in \mathbb{N})$  be a sequence with*

$$\lim_{n \rightarrow \infty} [\tfrac{1}{2} \kappa (1 - \epsilon) p \delta c_n - \log n] = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} c_n^2 / n = 0.$$

Suppose that

$$\frac{1}{n - \lfloor \epsilon n \rfloor} \sum_{v=\lfloor \epsilon n \rfloor + 1}^n \mathbb{1}\{|\underline{C}_{n,p}^\epsilon(v)| \geq 2c_n\} \geq \kappa \quad \text{with high probability.}$$

Then there exists a coupling of the networks  $(G_n^\epsilon(p))_n$  and  $(\underline{G}_n^\epsilon(p))_n$  such that  $\underline{G}_n^\epsilon(p) \leq G_n^\epsilon(p)$  and, with high probability, all connected components in  $\underline{G}_n^\epsilon(p)$  with at least  $2c_n$  vertices belong to one connected component in  $G_n^\epsilon(p)$ .

Lemma 4.11 in the case  $\epsilon = 0$  and  $p = 1$  is Proposition 4.1 in [15]. The proof is valid for  $\epsilon \in [0, 1)$ ,  $p \in (0, 1]$  up to obvious changes and is therefore omitted.

**Proof of Theorem 2.1.** Choose  $c_n = (\log n)^2$ . By Lemma 4.9, we have in probability

$$\limsup_{n \rightarrow \infty} \frac{|C_{n,p}^\epsilon|}{(1-\epsilon)n} \leq \limsup_{n \rightarrow \infty} \max \left\{ \frac{c_n}{(1-\epsilon)n}, M_{n,p}^\epsilon(c_n) \right\} \leq \zeta^\epsilon(p).$$

Moreover, for  $\delta \in (0, f(0))$ , another application of Lemma 4.9 implies that  $M_{n,p}^\epsilon(2c_n, f - \delta)$  converges to  $\zeta^\epsilon(p, f - \delta)$  in probability. Hence, Lemma 2.11 and Lemma 4.11 imply that for all  $\delta' > 0$ ,  $|C_{n,p}^\epsilon| \geq (n - \lfloor \epsilon n \rfloor)(\zeta^\epsilon(p) - \delta')$  with high probability. This concludes the proof.  $\square$

## 5 Variations and other models

We study the preferential attachment network with a non-linear attachment rule in Section 5.1, and inhomogeneous random graphs and the configuration model in Sections 5.2.2 and 5.2.1, respectively.

### 5.1 Proof of Theorem 1.4

Theorem 1.4 is an immediate consequence of Theorem 1.1 and a stochastic domination result on the level of the networks. We make use of the notation and terminology introduced in Section 1.4.

*Proof of Theorem 1.4.* First suppose that  $f$  is a L-class attachment rule with  $\bar{f} \geq f \geq \underline{f}$  for two affine attachment rules given by  $\bar{f}(k) = \gamma k + \beta_u$  and  $\underline{f}(k) = \gamma k + \beta_l$  for  $k \in \mathbb{N}_0$ . There exists a natural coupling of the networks generated by these attachment rules such that

$$\bar{G}_n \geq G_n \geq \underline{G}_n \quad \text{for all } n \in \mathbb{N}.$$

This ordering is retained after percolation and implies an ordering  $\bar{p}_c(\epsilon) \leq p_c(\epsilon) \leq \underline{p}_c(\epsilon)$  of the critical percolation parameters of the networks. Applying Theorem 1.1 to  $\bar{f}$  and  $\underline{f}$ , we obtain positive constants  $C_1, \dots, C_4$  such that, for small  $\epsilon \in (0, 1)$ ,

$$\frac{C_1}{\log(1/\epsilon)} \leq p_c(\epsilon) \leq \frac{C_2}{\log(1/\epsilon)} \quad \text{if } \gamma = \frac{1}{2}, \quad C_3 \epsilon^{\gamma-1/2} \leq p_c(\epsilon) \leq C_4 \epsilon^{\gamma-1/2} \quad \text{if } \gamma > \frac{1}{2},$$

and the result follows.

Now let  $f$  be a C-class attachment rule. Concavity of  $f$  implies that the increments  $\gamma_k := f(k+1) - f(k)$ , for  $k \in \mathbb{N}_0$ , form a non-increasing sequence converging to  $\gamma$ . In particular, with  $\underline{f}(k) := \gamma k + f(0)$ , we get  $f(k) = \sum_{l=0}^{k-1} \gamma_l + f(0) \geq \gamma k + f(0) = \underline{f}(k)$ , for all  $k \in \mathbb{N}_0$ . To obtain a corresponding upper bound, let  $\beta_j := f(j) - \gamma_j j$ . Then  $\beta_j > 0$  and for all  $k \in \mathbb{N}_0$ ,

$$f(k) - \beta_j = f(k) - f(j) + \gamma_j j = \begin{cases} -\sum_{l=k}^{j-1} \gamma_l + \gamma_j j \leq -(j-k)\gamma_j + \gamma_j j & \text{if } k \leq j \\ \sum_{l=j}^{k-1} \gamma_l + \gamma_j j \leq (k-j)\gamma_j + \gamma_j j & \text{if } k \geq j \end{cases} = \gamma_j k.$$

Hence, the attachment rule given by  $\bar{f}_j(k) := \gamma_j k + \beta_j$ , for  $k \in \mathbb{N}_0$ , satisfies  $\bar{f}_j \geq f$ , and we can use the same coupling as in the first part of the proof to obtain

$$\bar{p}_c^{(j)}(\epsilon) \leq p_c(\epsilon) \leq \underline{p}_c(\epsilon),$$

where  $\bar{p}_c^{(j)}(\epsilon)$  corresponds to the network with attachment rule  $\bar{f}_j$ . Since  $\gamma_j \downarrow \gamma$ , we have  $\gamma_j \in [\frac{1}{2}, 1)$  for large  $j$ . Theorem 1.1 yields, for  $\gamma > \frac{1}{2}$ , constants  $C, C_j > 0$  such that

$$\log C_j + (\gamma_j - 1/2) \log \epsilon \leq \log p_c(\epsilon) \leq \log C + (\gamma - 1/2) \log \epsilon.$$

Dividing by  $\log \epsilon$  and then taking first  $\epsilon \downarrow 0$  and then  $j \rightarrow \infty$  yields the claim for  $\gamma > \frac{1}{2}$ . In case  $\gamma = \frac{1}{2}$  it could happen that  $\gamma_j > \frac{1}{2}$  for all  $j \in \mathbb{N}$ . In this situation, Theorem 1.1 does not give a bound on the right scale. Therefore, we can use only the upper bound on  $p_c(\epsilon)$  which gives the stated result.  $\square$

### 5.2 Other models

In this section, we study vulnerability of two other classes of robust network models.

#### 5.2.1 Configuration model

The configuration model is a natural way to construct a network with given degree sequence. Its close connection to the uniformly chosen simple graph with given degree sequence is explained in Section 7.4 in [23]. Existence of a giant component in the configuration model has been studied by Molloy and Reed [30] and Janson and Luczak [26, 25]. Recall from Section 1.5.1 that we write  $D$  for the weak limit of the

degree of a uniformly chosen vertex. Janson and Luczak [26] showed that if (1.6) holds and  $\mathbb{P}(D = 2) < 1$ , then

$$(\mathbb{G}_n^{(\text{CM})} : n \in \mathbb{N}) \text{ has a giant component} \Leftrightarrow \mathbb{E}[D(D-1)] > \mathbb{E}D.$$

Janson [25] found a simple construction that allows to obtain a corresponding result for the network after random or deterministic removal of vertices (or edges), where the retention probability of a vertex can depend on its degree. Let  $\pi = (\pi_k)_{k \in \mathbb{N}}$  be a sequence of retention probabilities with  $\pi_k \mathbb{P}(D = k) > 0$  for some  $k$ . Every vertex  $i$  is removed with probability  $1 - \pi_{d_i}$  and kept with probability  $\pi_{d_i}$ , independently of all other vertices. Janson describes the network after percolation as follows [25, page 90]: For each vertex  $i$  replace it with probability  $1 - \pi_{d_i}$  by  $d_i$  new vertices of degree 1. Then construct the configuration model  $\mathbb{G}_n^{(\text{CM}), \pi}$  corresponding to the new degree sequence and larger number of vertices and remove from this graph uniformly at random vertices of degree 1 until the correct number of vertices for  $\mathbb{G}_n^{(\text{CM})}$  after percolation is reached. The removal of these surplus vertices cannot destroy or split the giant component since the vertices are of degree 1. Hence, it suffices to study the existence of nonexistence of a giant component in  $\mathbb{G}_n^{(\text{CM}), \pi}$ .

To construct  $\mathbb{G}_n^{(\text{CM}), \epsilon}(p)$ , we remove the  $\lfloor \epsilon n \rfloor$  vertices with the largest degree from  $\mathbb{G}_n^{(\text{CM})}$  and then run vertex percolation with retention probability  $p$  on the remaining graph. In general, this does not fit exactly into the setup of Janson. To emulate the behaviour, we denote by  $n_j$  the number of vertices with degree  $j$  in the graph and let  $K_n = \inf\{k \in \mathbb{N}_0 : \sum_{j=k+1}^{\infty} n_j \leq \lfloor \epsilon n \rfloor\}$ . Then all vertices with degree larger than  $K_n$  are deterministically removed in  $\mathbb{G}_n^{(\text{CM}), \epsilon}(p)$ , i.e.  $\pi_j = 0$  for  $j \geq K_n + 1$ . In addition, we deterministically remove  $\lfloor \epsilon n \rfloor - \sum_{j=K_n+1}^{\infty} n_j$  vertices of degree  $K_n$ , while all other vertices are subject to vertex percolation with retention probability  $p$ . In particular,  $\pi_j = p$  for  $j \leq K_n - 1$ . Denote by  $F(x) := \mathbb{P}(D \leq x)$ , for  $x \geq 0$ , and by  $[1 - F]^{-1}$  the generalized inverse of  $[1 - F]$ , that is  $[1 - F]^{-1}(u) = \inf\{k \in \mathbb{N}_0 : [1 - F](k) \leq u\}$  for all  $u \in (0, 1)$ . One easily checks that  $K_n \in \{m, m+1\}$  for all large  $n$ , where  $m := [1 - F]^{-1}(\epsilon)$ . Using this observation, it is not difficult to adapt Janson's proof to show that

$$(\mathbb{G}_n^{(\text{CM}), \epsilon}(p) : n \in \mathbb{N}) \text{ has a giant component} \Leftrightarrow p > p_c(\epsilon),$$

where

$$p_c(\epsilon) := \frac{\mathbb{E}D}{\mathbb{E}[D(D-1)\mathbb{1}_{D \leq m}] - m(m-1)(\epsilon - [1 - F](m))}.$$

*Proof of Theorem 1.6.* Let  $U$  be a uniformly distributed random variable on  $(0, 1)$ . Then  $[1 - F]^{-1}(U)$  has the same distribution as  $D$  and

$$\begin{aligned} \mathbb{E}[D(D-1)\mathbb{1}_{\{D \leq m\}}] - m(m-1)(\epsilon - [1 - F](m)) \\ = \mathbb{E}[[1 - F]^{-1}(U)([1 - F]^{-1}(U) - 1)\mathbb{1}_{\{U \geq \epsilon\}}] \asymp \mathbb{E}[[1 - F]^{-1}(U)^2\mathbb{1}_{\{U \geq \epsilon\}}]. \end{aligned}$$

The assumption  $[1 - F](k) \sim Ck^{-1/\gamma}$  as  $k \rightarrow \infty$  implies that  $[1 - F]^{-1}(u) \sim C^\gamma u^{-\gamma}$  as  $u \downarrow 0$ . Let  $u_0 > 0$  such that

$$\frac{1}{2} \leq \frac{[1 - F]^{-1}(u)}{C^\gamma u^{-\gamma}} \leq \frac{3}{2} \quad \text{for all } u \leq u_0.$$

Since  $[1 - F]^{-1}(u)^2$  is not integrable around zero, but bounded on  $[u_0, 1]$ , we deduce

$$\begin{aligned} \mathbb{E}[[1 - F]^{-1}(U)^2\mathbb{1}_{\{U \geq \epsilon\}}] &= \int_{\epsilon}^1 [1 - F]^{-1}(u)^2 du \asymp \int_{\epsilon}^{u_0} [1 - F]^{-1}(u)^2 du \\ &\asymp \int_{\epsilon}^{u_0} u^{-2\gamma} du \asymp \begin{cases} \log(1/\epsilon) & \text{if } \gamma = \frac{1}{2}, \\ \epsilon^{1-2\gamma} & \text{if } \gamma > \frac{1}{2}. \end{cases} \quad \square \end{aligned}$$

### 5.2.2 Inhomogeneous random graphs

The classical Erdős-Rényi random graph can be generalized by giving each vertex a weight and choosing the probability for an edge between two vertices as an increasing function of their weights. Suppose that  $\kappa: (0, 1] \times (0, 1] \rightarrow (0, \infty)$  is a symmetric, continuous kernel with

$$\int_0^1 \int_0^1 \kappa(x, y) dx dy < \infty \quad (5.1)$$

and recall from (1.7) that in the inhomogeneous random graph  $G_n^{(\kappa)}$ , the edge  $\{i, j\}$  is present with probability  $\frac{1}{n} \kappa(\frac{i}{n}, \frac{j}{n}) \wedge 1$ , independently of all other edges. We assume that vertices are ordered in decreasing order of privilege, i.e.  $\kappa$  is non-increasing in both components. Bollobás et al. show in Theorem 3.1 and Example 4.11 of [6] that

$$(G_n^{(\kappa), \epsilon}(p): n \in \mathbb{N}) \text{ has a giant component} \Leftrightarrow p > p_c(\epsilon) := \|T_\kappa\|_{L^2(\epsilon, 1)}^{-1}, \quad (5.2)$$

where

$$T_\kappa g(x) = \int_\epsilon^1 \kappa(x, y) g(y) dy, \quad \text{for all } x \in (\epsilon, 1),$$

for all measurable functions  $g$  such that the integral is well-defined, and  $\|\cdot\|_{L^2(\epsilon, 1)}$  denotes the operator norm on the  $L^2$ -space with respect to the Lebesgue measure on  $(\epsilon, 1)$ . The same result holds for a version of the Norros-Reittu model in which edges between different vertex pairs are independent and edge  $\{i, j\}$  is present with probability  $1 - e^{-\kappa(i/n, j/n)/n}$  for all  $i, j \in \{1, \dots, n\}$ . Consequently, the estimates given in Theorem 1.7 hold for this model, too.

*Proof of Theorem 1.7.* Since  $\kappa^{(\text{CL})}$  and  $\kappa^{(\text{PA})}$  are positive, symmetric, continuous kernels satisfying (5.1), the first part of the theorem follows immediately from (5.2). By definition

$$\|T_\kappa\|_{L^2(\epsilon, 1)} = \sup \{ \|T_\kappa g\|_{L^2(\epsilon, 1)} : \|g\|_{L^2(\epsilon, 1)} \leq 1 \}.$$

For a rank one kernel  $\kappa(x, y) = \psi(x)\psi(y)$ , the operator norm of  $T_\kappa$  is attained at the function  $\psi/\|\psi\|_{L^2(\epsilon, 1)}$  with  $\|T_\kappa\|_{L^2(\epsilon, 1)} = \|\psi\|_{L^2(\epsilon, 1)}^2$ . Hence,

$$\|T_{\kappa^{(\text{CL})}}\|_{L^2(\epsilon, 1)} = \int_\epsilon^1 x^{-2\gamma} dx = \begin{cases} \log(1/\epsilon) & \text{if } \gamma = 1/2, \\ \frac{1}{1-2\gamma} [1 - \epsilon^{1-2\gamma}] & \text{if } \gamma \neq \frac{1}{2}, \end{cases}$$

and

$$p_c^{(\text{CL})}(\epsilon) = \begin{cases} (1-2\gamma) \frac{1}{1-\epsilon^{1-2\gamma}} & \text{if } \gamma < \frac{1}{2}, \\ \frac{1}{\log(1/\epsilon)} & \text{if } \gamma = \frac{1}{2}, \\ (2\gamma-1) \epsilon^{2\gamma-1} \frac{1}{1-\epsilon^{2\gamma-1}} & \text{if } \gamma > \frac{1}{2}. \end{cases}$$

Now suppose that  $\gamma > 1/2$ . By Cauchy-Schwarz's inequality and the symmetry of  $\kappa^{(\text{PA})}$ ,

$$\begin{aligned} \|T_{\kappa^{(\text{PA})}}\|_{L^2(\epsilon, 1)}^2 &\leq \int_\epsilon^1 \int_\epsilon^1 \kappa^{(\text{PA})}(x, y)^2 dy dx = 2 \int_\epsilon^1 \int_\epsilon^x x^{2(\gamma-1)} y^{-2\gamma} dy dx \\ &\leq 2 \int_0^1 x^{2(\gamma-1)} dx \int_\epsilon^\infty y^{-2\gamma} dy = \frac{2}{(2\gamma-1)^2} \epsilon^{1-2\gamma}. \end{aligned}$$

For the lower bound, let  $c_\epsilon = \sqrt{2\gamma - 1}\epsilon^{\gamma-1/2}$  and  $g(x) = c_\epsilon x^{-\gamma}$ . Then  $\|g\|_{L^2(\epsilon,1)} \leq 1$  and

$$\begin{aligned} \|T_{\kappa^{(\text{PA})}}\|_{L^2(\epsilon,1)}^2 &\geq \|T_{\kappa^{(\text{PA})}}g\|_{L^2(\epsilon,1)}^2 \geq \int_\epsilon^1 \left( \int_\epsilon^x \kappa^{(\text{PA})}(x,y)g(y)dy \right)^2 dx \\ &= \frac{c_\epsilon^2}{(2\gamma-1)^2} \int_\epsilon^1 x^{2(\gamma-1)} [\epsilon^{1-2\gamma} - x^{1-2\gamma}]^2 dx \\ &\geq \frac{c_\epsilon^2}{(2\gamma-1)^2} \int_\epsilon^1 x^{2(\gamma-1)} [\epsilon^{2(1-2\gamma)} - 2\epsilon^{1-2\gamma}x^{1-2\gamma}] dx \\ &= \frac{\epsilon^{1-2\gamma}}{(2\gamma-1)^2} \left[ 1 - \epsilon^{2\gamma-1} + 2(2\gamma-1)\epsilon^{2\gamma-1} \log \epsilon \right]. \end{aligned}$$

The claim follows.  $\square$

## References

- [1] Réka Albert, István Albert, and Gary L. Nakarado, *Structural vulnerability of the North American power grid*, Phys. Rev. E **69** (2004), 025103.
- [2] Réka Albert, Hawoong Jeong, and Albert-László Barabási, *Error and attack tolerance of complex networks*, Nature **406** (2000), 378–382.
- [3] Albert-László Barabási and Réka Albert, *Emergence of Scaling in Random Networks*, Science **286** (1999), no. 5439, 509–512. MR-2091634
- [4] Shankar Bhamidi, Remco van der Hofstad, and Gerard Hooghiemstra, *Extreme value theory, Poisson-Dirichlet distributions, and first passage percolation on random networks*, Adv. Appl. Probab. **42** (2010), no. 3, 706–738. MR-2779556
- [5] John D. Biggins and Andreas E. Kyprianou, *Measure change in multitype branching*, Adv. Appl. Probab. **36** (2004), no. 2, 544–581. MR-2058149
- [6] Béla Bollobás, Svante Janson, and Oliver Riordan, *The Phase Transition in Inhomogeneous Random Graphs*, Random Structures Algorithms **31** (2007), no. 1, 3–122. MR-2337396
- [7] Béla Bollobás and Oliver Riordan, *Robustness and Vulnerability of Scale-Free Random Graphs*, Internet Math. **1** (2003), no. 1, 1–35. MR-2076725
- [8] Duncan S. Callaway, Mark E. J. Newman, Steven H. Strogatz, and Duncan J. Watts, *Network Robustness and Fragility: Percolation on Random Graphs*, Phys. Rev. Lett. **85** (2000), no. 25, 5468–5471.
- [9] Fan Chung and Linyuan Lu, *The average distances in random graphs with given expected degrees*, Proc. Nat. Acad. Sci. USA **99** (2002), no. 25, 15879–15882. MR-1944974
- [10] ———, *Complex graphs and networks*, Regional Conference Series in Mathematics, vol. 107, Co-publication of the AMS and CBMS, 2006.
- [11] Reuven Cohen, Keren Erez, Daniel ben-Avraham, and Shlomo Havlin, *Breakdown of the Internet under Intentional Attack*, Phys. Rev. Lett. **86** (2001), 3682–3685.
- [12] ———, *Reply*, Phys. Rev. Lett. **87** (2001), 219802.
- [13] Steffen Dereich, Christian Mönch, and Peter Mörters, *Typical distances in ultrasmall random networks*, Adv. Appl. Probab. **44** (2012), 583–601. MR-2977409
- [14] Steffen Dereich and Peter Mörters, *Random networks with sublinear preferential attachment: degree evolutions*, Electron. J. Probab. **14** (2009), 1222–1267. MR-2511283
- [15] ———, *Random networks with sublinear preferential attachment: the giant component*, Ann. Probab. **41** (2013), no. 1, 329–384. MR-3059201
- [16] Sander Dommers, Remco van der Hofstad, and Gerard Hooghiemstra, *Diameters in preferential attachment models*, J. Stat. Phys. **139** (2010), no. 1, 72–107. MR-2602984
- [17] Sergey N. Dorogovtsev and José F. F. Mendes, *Comment on “Breakdown of the Internet under Intentional Attack”*, Phys. Rev. Lett. **87** (2001), 219801.
- [18] Henri van den Esker, Remco van der Hofstad, Gerard Hooghiemstra, and Dmitri Znamenski, *Distances in random graphs with infinite mean degrees*, Extremes **8** (2006), no. 3, 111–141. MR-2275914

- [19] William Feller, *An introduction to probability theory and its applications*, vol. II, Wiley, New York, 1971. MR-0270403
- [20] Simon C. Harris, Marion Hesse, and Andreas E. Kyprianou, *Branching Brownian motion in a strip: survival near criticality*, Preprint arXiv:arXiv:1212.1444, 2012.
- [21] Theodore E. Harris, *The theory of branching processes*, Grundlehren Math. Wiss., Springer, Berlin, 1963. MR-0163361
- [22] Harro Heuser, *Functional analysis*, Wiley, Chichester, 1982. MR-0640429
- [23] Remco van der Hofstad, *Random graphs and complex networks*, Draft book, available online at <http://www.win.tue.nl/~rhofstad/NotesRGCN.html>, 2013.
- [24] Petter Holme, Beom Jun Kim, Chang No Yoon, and Seung Kee Han, *Attack vulnerability of complex networks*, Phys. Rev. E **65** (2002), 056109.
- [25] Svante Janson, *On percolation in random graphs with given vertex degrees*, Electron. J. Probab. **14** (2009), no. 5, 87–118. MR-2471661
- [26] Svante Janson and Malwina J. Luczak, *A New Approach to the Giant Component Problem*, Random Structures Algorithms **34** (2008), no. 2, 197–216. MR-2490288
- [27] Tosio Kato, *Perturbation Theory for Linear Operators*, Classics Math., Springer, Berlin, 1995, Reprint of the 1980 edition. MR-1335452
- [28] Peter D. Miller, *Applied asymptotic analysis*, Graduate Studies in Mathematics, vol. 75, Amer. Math. Soc., 2006. MR-2238098
- [29] Igor Mishkovski, Mario Biey, and Ljupco Kocarev, *Vulnerability of complex networks*, Commun. Nonlinear Sci. Numer. Simul. **16** (2011), no. 1, 341–349. MR-2679185
- [30] Michael Molloy and Bruce Reed, *A critical point for random graphs with a given degree sequence*, Random Structures Algorithms **6** (1995), no. 2-3, 161–180. MR-1370952
- [31] Christian Mönch, *Distances in preferential attachment networks*, Ph.D. thesis, University of Bath, 2013.
- [32] Ross G. Pinsky, *Positive Harmonic Functions and Diffusion*, Cambridge Stud. Adv. Math, vol. 45, Cambridge University Press, Cambridge, 1995. MR-1326606

**Acknowledgments.** We thank Steffen Dereich for useful discussions and an anonymous referee for their very careful reading of the paper and numerous valuable suggestions.



---

# Electronic Journal of Probability

## Electronic Communications in Probability

---

### Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

### Economical model of EJP-ECP

- Low cost, based on free software (OJS<sup>1</sup>)
- Non profit, sponsored by IMS<sup>2</sup>, BS<sup>3</sup>, PKP<sup>4</sup>
- Purely electronic and secure (LOCKSS<sup>5</sup>)

### Help keep the journal free and vigorous

- Donate to the IMS open access fund<sup>6</sup> (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

---

<sup>1</sup>OJS: Open Journal Systems <http://pkp.sfu.ca/ojs/>

<sup>2</sup>IMS: Institute of Mathematical Statistics <http://www.imstat.org/>

<sup>3</sup>BS: Bernoulli Society <http://www.bernoulli-society.org/>

<sup>4</sup>PK: Public Knowledge Project <http://pkp.sfu.ca/>

<sup>5</sup>LOCKSS: Lots of Copies Keep Stuff Safe <http://www.lockss.org/>

<sup>6</sup>IMS Open Access Fund: <http://www.imstat.org/publications/open.htm>